

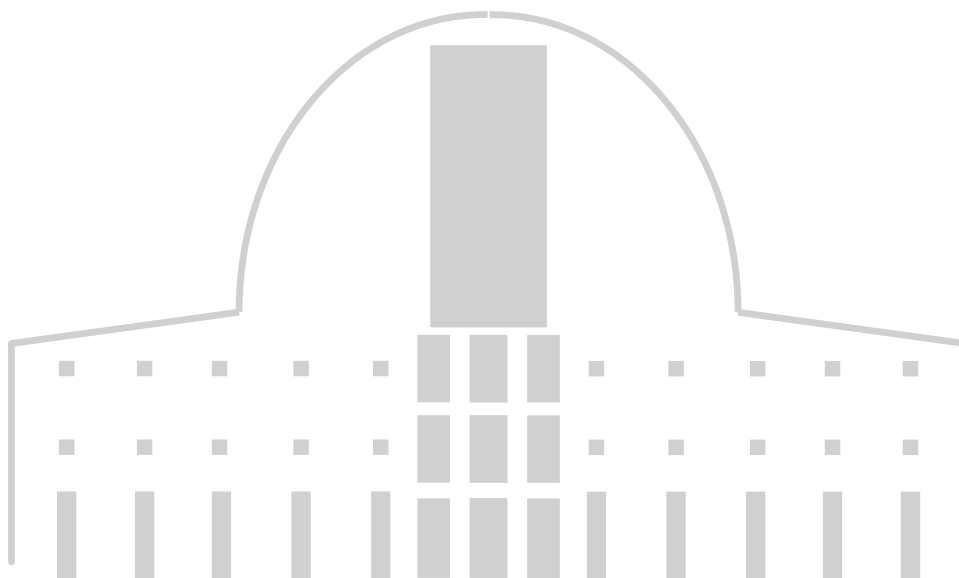
Post-transcriptional regulation mediated by 3'-UTRs in bacteria

Tesis Doctoral para optar al grado de Doctor, presentada por:

Igor Ruiz de los Mozos Aliaga

Directores: **Dr. Alejandro Toledo Arana - Prof. Íñigo Lasa Uzcudun**

Pamplona, 09 de junio de 2014



Universidad Pública de Navarra
Departamento de Producción Agraria
Instituto de Agrobiotecnología

Prof. ÍÑIGO LASA UZCUDUN, Catedrático de Microbiología del Departamento de Producción Agraria de la Universidad Pública de Navarra,

Dr. ALEJANDRO TOLEDO ARANA, Científico Titular del Consejo Superior de Investigaciones Científicas en el Instituto de Agrobiotecnología,

INFORMAN:

Que la presente memoria de Tesis Doctoral titulada “Post-transcriptional regulation mediated by 3’-UTRs in bacteria” y elaborada por IGOR RUIZ DE LOS MOZOS ALIAGA ha sido realizada bajo su dirección y que cumple las condiciones exigidas por la legislación vigente para optar al grado de Doctor.

Y para que así conste, firman la presente en Pamplona, a 09 de junio de 2014.

Fdo. Íñigo Lasa Uzcudun

Fdo. Alejandro Toledo Arana

"This is not the end. It is not even the beginning of the end.

But it is, perhaps, the end of the beginning."

Winston Churchill

A mis padres

AGRADECIMIENTOS

Quisiera transmitir mi mas sincero agradecimiento a todas las personas e instituciones que de un modo u otro han hecho posible la realización de este trabajo. A la Universidad Pública de Navarra, al Departamento de Producción Agraria y al Instituto de Agrobiotecnología.

Me gustaría comenzar agradeciendo a Rosa, Jon y Mari José, que me ayudaron y alentaron en mis comienzos en investigación.

A Iñigo, por formar un ambiente de trabajo exigente pero distendido, en el que existe un compañerismo sin igual. Por ser una gran persona y por preocuparse tanto de mi vida personal como del resultado del último experimento.

A Alejandro, por enseñarme, dirigirme y ayudarme todos y cada uno de los días que he estado aquí; por no hacerme pagar las cervezas que nos hemos jugado ante el resultado de algún experimento y porque sin él esta tesis no habría sido posible.

A todos los Biofilms, Solano, Jaio, Bego, Nekane, Marta, Irene, Gabriel, Violeta, Villanueva, Saio, Tana, Carlos, Sonia, Juani y Amaya; por vuestro cariño y amistad; por el compañerismo que me habéis demostrado y por enseñarme y aconsejarme en el día a día. No puedo reivindicar todas las ideas aquí contenidas porque muchas han salido de las pautas que me habéis transmitido.

A Latasa, Kike y Echeverz, por brindarme vuestra amistad, consejo y consuelo en los cientos de comidas, cafés y pitillos que hemos compartido en esta larga tesis.

A todos los compañeros del Instituto de Agrobiotecnología, con mención especial a los BAZs por propiciar “buen rollo” en la primera planta; al grupo de Juncal por demostrarme que soy uno más de ellos; a Víctor y su inmejorable predisposición; y finalmente al servicio de apoyo a la investigación, que siempre están dispuestos a ayudar. Gracias por compartir instalaciones, cafés, almuerzos y buenos ratos fuera y dentro del Instituto.

A Cecilia Arraiano y todos los integrantes de su laboratorio, en especial a Margarida y Susana por acogerme en mi estancia doctoral y por enseñarme como funciona el mundo de la ciencia y el RNA.

Al laboratorio de Pascale Romby, por demostrar una gran implicación y profesionalidad en la colaboración que mantenemos.

A Ido, por aguantar tantas charlas de mis truños e inquietudes, por quererme incondicionalmente y por dar rienda suelta a mis sueños. Porque nadie mas que tú ha sufrido con esta tesis, gracias.

A toda la cuadrilla de Zulueta, porque me demostráis vuestra amistad cada vez que nos vemos; porque me ayudáis a desconectar y tirar para adelante.

A mis padres, por confiar en mí hasta el último momento y por apoyarme en mis difíciles decisiones. Sin vosotros no estaría aquí, sabéis que os quiero y que sois lo mas importante para mí.

Sin todos vosotros esto no habría sido lo mismo. ¡¡Gracias a todos!!

Para la realización de este trabajo el doctorando ha disfrutado de una beca para la Formación de Personal Investigador (FPI) asociada al proyecto BIO2008-05284-C02-01 del Ministerio de Educación y Ciencia.

Otros proyectos de investigación que han contribuido en la realización de este trabajo han sido:

- A) Functional genomic characterization of molecular determinants for staphylococcal fitness, virulence and drug resistance, VI Programa Marco UE (LSHM-CT-2006, 019064).
- B) Global analysis of antisense regulatory mechanisms in *Staphylococcus aureus*, ERANET Pathogenomics (PIM2010EPA-00606).
- C) Análisis de la regulación post-transcripcional mediada por proteínas de unión a RNA en *Staphylococcus aureus*. Ministerio de Economía y Competitividad (BFU2011-23222).

ÍNDICE

| | |
|---|----|
| RESUMEN | 3 |
| SUMMARY | 7 |
| 1. INTRODUCCIÓN | 11 |
| 1.1. EL TRANSCRIPTOMA | 11 |
| 1.1.1. Métodos para la determinación de los mapas transcriptómicos | 13 |
| 1.2. LA REGULACIÓN POST-TRANSCRIPCIONAL..... | 18 |
| 1.2.1. La degradación del RNA..... | 19 |
| 1.2.2. Regulación mediada por RNAs | 23 |
| 1.2.3. RNAs reguladores pequeños (sRNAs) | 23 |
| 1.2.4. RNAs antisentidos | 24 |
| 1.2.5. Regiones no traducidas (UTRs) del mRNA | 28 |
| 1.3. <i>STAPHYLOCOCCUS AUREUS</i> COMO MODELO DE ESTUDIO | 36 |
| 1.3.1. Características de <i>S. aureus</i> | 37 |
| 1.3.2. Patogenia de <i>S. aureus</i> y multi-resistencia antibiótica | 38 |
| 1.3.3. La capacidad para formar biofilms de <i>S. aureus</i> | 40 |
| 1.3.3.1. El proceso de formación de biofilms en <i>S. aureus</i> | 41 |
| 1.3.3.2. La matriz del biofilm en <i>S. aureus</i> | 42 |
| 1.3.3.3. La regulación de la síntesis del exopolisacárido PIA-PNAG | 45 |
| 2. OBJETIVOS DE LA TESIS DOCTORAL | 49 |
| 3. MATERIALES Y MÉTODOS | 53 |
| 3.1. Oligonucleotides, plasmids, bacterial strains and culture conditions | 53 |
| 3.2. RNA extractions | 53 |
| 3.3. RNA sequencing | 54 |
| 3.4. cDNA synthesis, labelling and tiling array hybridization..... | 55 |
| 3.5. Microarray data analysis | 56 |
| 3.6. cDNA libraries for RNA-seq, read mapping and statistics analysis | 57 |
| 3.7. Simultaneous mapping of 5'- and 3'-ends of RNA molecules | 57 |
| 3.8. Riboprobes synthesis..... | 58 |
| 3.9. Northern blots | 59 |
| 3.10. Chromosomal allelic exchange | 60 |
| 3.11. Plasmid constructions | 60 |
| 3.12. Quantitative reverse transcription PCR..... | 61 |
| 3.13. Western blot analysis..... | 62 |
| 3.14. mRNA stability assays | 63 |
| 3.15. mRNA dimerization assays on native agarose gel electrophoresis | 64 |
| 3.16. Gel shift assays..... | 65 |
| 3.17. Purification of recombinant <i>S. aureus</i> RNase III | 66 |

| | |
|---|------------|
| 3.18. In vitro RNase III activity assays | 67 |
| 3.19. Toeprinting assays | 68 |
| 3.20. PIA-PNAG quantification | 69 |
| 3.21. Biofilm formation assay | 70 |
| 3.22. RNA structural probing | 71 |
| 4. RESULTADOS | 79 |
| 4.1. Identification of long 3'-UTRs in the <i>S. aureus</i> transcriptome | 79 |
| 4.2. 3'-UTRs can be predicted in other bacteria | 83 |
| 4.3. The <i>icaR</i> mRNA contains a highly conserved long 3'-UTR | 88 |
| 4.4. 3'-UTR <i>icaR</i> post-transcriptionally modulates IcaR expression | 93 |
| 4.5. Base pairing interaction between 3'- and 5'-UTR of <i>icaR</i> mRNA | 97 |
| 4.6. The 3'-5'-UTRs pairing provides a substrate for RNase III cleavage | 102 |
| 4.7. Interaction between <i>icaR</i> 3'-UTR and SD regions prevents the formation of the translational initiation complex | 104 |
| 4.8. Biological relevance of <i>icaR</i> mRNA 3'-UTR regulation | 108 |
| 4.9. Temperature shifts modulate IcaR expression and biofilm formation | 112 |
| 4.10. Temperature controls the interaction between 5' and 3'-UTRs of <i>icaR</i> mRNA ... | 115 |
| 4.11. <i>icaR</i> 5'-UTR forms a three way-helix RNA structure | 117 |
| 4.12. The three way-helical junction is disrupted by the binding of 3'-UTR in a temperature dependent manner | 120 |
| 5. DISCUSIÓN | 127 |
| 5.1. Regulación de la traducción mediante la interacción entre 3'- y 5'-UTRs | 127 |
| 5.2. Relevancia biológica de la interacción entre la 3'-UTR y la 5'-UTR | 131 |
| 5.3. La temperatura modula la interacción entre la 3'-UTR y la 5'-UTR del mRNA de <i>icaR</i> | 135 |
| 5.4. Los elementos reguladores localizados en las 3'-UTRs estarían presentes en todas las bacterias | 139 |
| 6. CONCLUSIONES | 145 |
| 7. REFERENCIAS | 150 |

RESUMEN

RESUMEN

En organismos eucariotas, la presencia de elementos reguladores en las regiones 3' no traducidas (3'-UTRs) del RNA mensajero (mRNA), que controlan su estabilidad y la eficiencia de su traducción, está ampliamente reconocida. En cambio, la posible relevancia de las 3'-UTRs como elementos funcionales de los mRNAs bacterianos ha sido menospreciada. En esta Tesis Doctoral, se presentan evidencias que muestran que un tercio de los mRNAs de *Staphylococcus aureus*, uno de los patógenos más importantes a nivel mundial, contienen 3'-UTRs mayores de 100 nucleótidos (nt) de largo y, por lo tanto, con capacidad para albergar funciones reguladoras. Además, se vio que la mayoría de estas 3'-UTRs incluyen un terminador de transcripción Rho-independiente. Basados en esta información, fue posible predecir las 3'-UTRs en otras bacterias. Así, mediante el análisis de 25 genomas, se muestra que las 3'-UTRs largas están ampliamente distribuidas en el mundo procariota. Con el fin de evaluar el rol que las 3'-UTRs pueden ejercer en el control de la expresión de los mRNAs, se seleccionó la 3'-UTR larga del mRNA de *icaR*, que codifica para el represor de la síntesis de principal exopolisacárido de la matriz del biofilm de *S. aureus*. Se descubrió que una región de esta 3'-UTR hibrida con una región de la 5'-UTR que incluye la secuencia Shine-Dalgarno. Como resultado de esta interacción, se crea una región de RNA de cadena doble que bloquea la formación del complejo de iniciación de la traducción y que, además, es reconocida por la endoribonucleasa RNase III

que digiere este dúplex de RNA, promoviendo la degradación del mRNA de *icaR*. Además, se descubrió que cambios conformaciones en la estructura de la 5'-UTR controlan la interacción con la 3'-UTR en respuesta a la temperatura. A 23°C (temperatura ambiental), una región de la 5'-UTR hibrida con una zona de la región codificante (ORF) generando una estructura secundaria cerrada de tres horquillas que impide la interacción con la 3'-UTR, lo que causa la acumulación de la proteína IcaR y, en consecuencia, la inhibición de la formación del biofilm. En cambio, a 37°C (temperatura del hospedador) esta conformación estructural se abre permitiendo la interacción entre la 5'- y la 3'-UTR, inhibiendo la traducción de IcaR y facilitando la formación de biofilm. En resumen, esta Tesis Doctoral muestra un nuevo mecanismo de regulación post-transcripcional que modula la expresión génica en respuesta a cambios en la temperatura ambiental, a través de la modificación de la interacción entre dominios de RNA codificados tanto en la 3'-UTR como en la 5'-UTR de una misma molécula de mRNA.

Secciones de esta Tesis Doctoral han sido publicadas en:

2011. PNAS, 108 (50): 20172-20177

I. Lasa, A. Toledo-Arana, A. Dobin, M. Villanueva, I. Ruiz de los Mozos, M. Vergara-Irigaray, V. Segura, D. Fagegaltier, J. Penadés, J. Valle, C. Solano, T. Gingeras. Genome-wide antisense transcription drives mRNA processing in bacteria.

2013. PLOS Genetics, 9(12): e1004001

I. Ruiz de los Mozos, M. Vergara-Irigaray, V. Segura, M. Villanueva, N. Bitarte, M. Saramago, S. Domingues, C. M. Arraiano, P. Fechter, P. Romby, J. Valle, C. Solano, I. Lasa and A. Toledo-Arana. Base Pairing Interaction Between 5'- and 3'-UTRs Controls *icaR* mRNA Translation in *Staphylococcus aureus*.

SUMMARY

SUMMARY

The presence of regulatory elements in the 3' untranslated region (3'-UTR) of eukaryotic mRNAs controlling RNA stability and translation efficiency is widely recognized. In contrast, the relevance of 3'-UTRs in bacterial mRNA functionality has been disregarded. Here, we report evidences showing that around one-third of the mapped mRNAs of the major human pathogen *Staphylococcus aureus* carry 3'-UTRs longer than 100-nt and thus, potential regulatory functions. We also found that most of the long 3'-UTR ends in a Rho-independent transcriptional terminator. Based on this information, it is possible to predict 3'-UTRs in any bacteria. Thus, we analysed 25 genomes and found that 3'-UTRs longer than 100-nt are broadly distributed in prokaryotes. To evaluate the role that 3'-UTRs may play in controlling mRNA expression, we selected the long 3'-UTR of *icaR* mRNA, which encodes the repressor of the main exopolysaccharidic compound of the *S. aureus* biofilm matrix. We showed that base pairing between the 3'-UTR and the Shine-Dalgarno (SD) region of *icaR* mRNA interferes with the translation initiation complex and generates a double-stranded substrate for RNase III. We also unveiled that the *icaR* 5'-UTR controls the 5'-3'-UTRs interaction in response to temperature. At environmental temperature (23°C), a three way-helical junction structure, generated by pairing of internal sequences from 5'-UTR and ORF regions, impairs the 5'-3'-UTR interaction, causing the accumulation of IcaR repressor and the inhibition of biofilm development. In contrast, at the human body temperature (37°C),

this structural conformation opens allowing the interaction of the 5'- and 3'-UTRs that inhibited IcaR translation leading to biofilm production. Our findings provide a singular example of a new potential post-transcriptional regulatory mechanism to modulate bacterial gene expression in response to temperature shifts through the interaction of a 3'-UTR with the 5'-UTR of the same mRNA.

Sections of this Doctoral Thesis has been published at:

2011. PNAS, 108 (50): 20172-20177.

I. Lasa, A. Toledo-Arana, A. Dobin, M. Villanueva, I. Ruiz de los Mozos, M. Vergara-Irigaray, V. Segura, D. Fagegaltier, J. Penadés, J. Valle, C. Solano, T. Gingeras. Genome-wide antisense transcription drives mRNA processing in bacteria.

2013. PLOS Genetics, 9(12): e1004001

I. Ruiz de los Mozos, M. Vergara-Irigaray, V. Segura, M. Villanueva, N. Bitarte, M. Saramago, S. Domingues, C. M. Arraiano, P. Fechter, P. Romby, J. Valle, C. Solano, I. Lasa and A. Toledo-Arana. Base Pairing Interaction Between 5'- and 3'-UTRs Controls *icaR* mRNA Translation in *Staphylococcus aureus*.

INTRODUCCIÓN

1. INTRODUCCIÓN

1.1. EL TRANSCRIPTOMA

El genoma contiene la información necesaria para realizar todas las funciones biológicas de un organismo. Para que estas instrucciones se materialicen, los genes deben de ser copiados a moléculas de ácido ribonucleico (RNA) conocidas como transcritos. El conjunto de transcritos presentes en una célula en un determinado momento se denomina transcriptoma. El análisis del transcriptoma comprende la integración de la información de las posiciones de inicio y final de cada transcrito, junto con sus niveles de expresión. Además, puede incluir los sitios de procesamiento generados por determinadas ribonucleasas. Toda esta información permite la creación de mapas transcriptómicos que dan sentido a las regiones cromosómicas mostrando la actividad transcripcional de los genes anotados en un determinado genoma, y además identifica aquellas regiones activas transcripcionalmente que no han podido ser predichas basados en los marcos de lectura abiertos (ORFs).

El transcriptoma de cualquier organismo está constituido por distintas clases de RNAs que son necesarios para llevar a cabo el dogma central de la biología, donde el RNA mensajero (mRNA) producido a partir de un gen, se presenta a los ribosomas que traducirán su secuencia de nucleótidos a una secuencia de aminoácidos que constituirá la proteína codificada por ese gen (Woese, 1967; Crick, 1970). Proporcionalmente, la gran mayoría corresponde a los *RNA ribosomales* (rRNAs, el ~80% del RNA total)

(Condon, 2007), que son los elementos estructurales básicos de los ribosomas, y los *RNAs de transferencias* (tRNAs) (~14 %) encargados de transportar los aminoácidos hasta el ribosoma durante el proceso de la traducción (Woese, 1967; Woese, 1987). Sólo una pequeña parte corresponde a los mRNAs que codifican para la producción de proteínas. Además, una parte muy importante y que ha pasado desapercibida hasta hace unos pocos años, corresponde a una clase de RNAs que no codifican para proteínas y que pueden ser tan numerosos como los propios mRNAs: los *RNAs reguladores*. Estos RNAs incluyen los *pequeños RNAs* (sRNAs) y los *RNAs antisentidos* (asRNAs), los cuales han sido implicados en la regulación de casi la totalidad de los procesos biológicos llevados a cabo durante la vida de un organismo y por lo tanto son esenciales para su desarrollo y supervivencia (Masse, 2002; Romby and Charpentier, 2009; Fröhlich and Jörg Vogel, 2009; Gottesman and Storz, 2011).

A diferencia del genoma, que se caracteriza por ser relativamente estable a lo largo de la vida del organismo en cuestión, el transcriptoma puede cambiar rápidamente dependiendo de las condiciones medioambientales. Pero además, es importante comprender que dentro de un mismo organismo, o una colonia de microorganismos (si es el caso de unicelulares), el transcriptoma puede ser diferente para cada unidad celular, aun estando en el mismo ambiente y siendo genéticamente iguales (van Vliet and Wren, 2009; Morozova *et al.*, 2009; Toledo-Arana and Solano, 2010; Croucher and Thomson, 2010). Por ello, descifrar el transcriptoma no es una tarea sencilla, pero es esencial si queremos determinar qué parte de

todo el genoma es importante en una determinada condición ambiental o proceso de desarrollo para así poder comprender cómo un organismo consigue adaptarse a los continuos cambios de su entorno y perpetuar su especie. De hecho la identificación del transcriptoma es tan, o incluso más, importante que la propia secuencia genómica.

Uno de los proyectos de secuenciación de transcriptomas más ambiciosos que ha agrupado a cientos de grupos de investigación, el Proyecto ENCODE, ha revelado que el genoma humano es transcrito en más de un 60% (Bernstein *et al.*, 2012). De este porcentaje sólo una pequeña parte corresponde a regiones cromosómicas que codifican para proteínas, el resto proviene de regiones sin función conocida. Esto ha resuelto una de la mayores paradojas de la evolución, ya que lo que hasta hace poco se consideraba “basura genómica” han resultado contener regiones con alta actividad transcripcional. Como resultado, y en contra de lo que asumíamos hasta ahora, la gran mayoría del genoma humano se transcribe en moléculas que no codifican para proteínas. A pesar de que su funciones no están claras para la gran mayoría de ellas, se ha demostrado que algunas cumplen funciones esenciales como por ejemplo en la diferenciación celular.

1.1.1. Métodos para la determinación de los mapas transcriptómicos

El análisis de los transcriptomas no sólo se limita a obtener un catálogo de los niveles de expresión de todas las especies de transcritos, también

permite determinar la estructura de los genes, delimitando con precisión donde se encuentra el inicio (extremos 5') y el final (extremos 3') de la transcripción. Conociendo cada inicio de transcripción, es posible la identificación de las regiones promotoras, que son esenciales para poder estudiar los procesos de regulación transcripcional. Al mismo tiempo, determinar cuáles son los extremos de los transcritos nos permite conocer las regiones no traducidas de un mRNA. Esto puede ayudar a descubrir tanto los eventos de regulación post-transcripcional que controlan la expresión de dicho transcrito como conocer el tipo de terminación de la transcripción que usa la RNA polimerasa para un transcrito en particular.

La primera técnica que se utilizó para analizar los transcritos de la levadura *Sacharomyces cerevisiae* fue el análisis en serie de la expresión génica (SAGE) y que además sirvió para acuñar el término *transcriptoma* (Velculescu *et al.*, 1997). El SAGE consistía en retrotranscribir los mRNAs para posteriormente fragmentar los cDNAs en pequeños *tags* de 10-14 nt. Estos *tags* se concatenaban en serie mediante la unión de adaptadores en ambos extremos, y se clonaban en un vector bacteriano. Posteriormente se secuenciaba este vector y se separaban los *tags* en base a la secuencia de los adaptadores. Así estos *tags* constituían una pequeña muestra de los transcritos presentes en esa situación concreta.

La siguiente aproximación técnica se realizó mediante biochips de ADN, ahora conocidos como microarrays (Schena *et al.*, 1995). Los microarrays son superficies sólidas donde se adhieren sondas de DNA. Estas sondas

detectan los cDNAs complementarios (que se han producido por retrotranscripción a partir del RNA objeto de estudio) conjugados con un marcador fluorescente.

Los microarrays permiten comparar de manera simultánea la expresión de cientos de genes en respuesta a una condición particular o tratamiento (Schena *et al.*, 1995; DeRisi *et al.*, 1996; Lashkari *et al.*, 1997). En las primeras versiones de los arrays, estos se hibridaban con los dos juegos cDNAs a comparar, los cuales estaban marcados de forma diferencial, generalmente con fluorocromos rojos y verdes. Como la hibridación de ambos cDNAs se realizaba en las mismas condiciones sobre el mismo array, las posibles variaciones entre experimentos se reducían mucho. Los primeros análisis a escala global utilizando esta tecnología se realizaron en levaduras para analizar los cambios en los perfiles de expresión génica cuando se añadía glucosa o galactosa al medio de crecimiento (Lashkari *et al.*, 1997). Una limitación de estos primeros arrays era que contenían una sola sonda complementaria para cada una de las ORFs, y en consecuencia sólo permitían cuantificar los niveles de los mRNAs correspondientes a dichas ORFs. Al utilizar una sola sonda por ORF, los resultados a veces eran variables si se comparaban con técnicas más precisas como la RT-PCR cuantitativa.

A pesar de estas limitaciones, el potencial de la técnica era evidente, por lo que varias empresas comenzaron a comercializarlos permitiendo su estandarización y la consiguiente mejora en la reproducibilidad. Además,

aumentaron el número de sondas que hibridaba sobre un gen, mejorando significativamente la robustez de los resultados porque todas las sondas de un determinado gen debían variar en la misma dirección para considerar que su expresión había cambiado entre las dos condiciones objeto de análisis. Sin embargo, estos arrays solo tenían sondas complementarias a las zonas que codifican para proteínas y algunos sRNA ya conocidos, pero no tenían sondas en las regiones intergénicas o en las cadenas antisentido y por tanto no permitían detectar la existencia de transcritos en zonas distintas a los genes ya definidos. Este problema fue solventado gracias a la mejora en las técnicas de litografía, que permitía fijar millones de sondas en un espacio mucho más reducido, y de la resolución de los escáneres, que distinguían la señal de cada una de estas sondas, ahora fijadas a unos pocos micrómetros de distancia. Esta técnica permitió el desarrollo de los nuevos microarrays de alta densidad conocidos popularmente como tiling arrays. Estos arrays poseen sondas solapadas que cubren ambas cadenas del genoma permitiendo mapear los transcritos sin el sesgo de la anotación genómica, con una resolución de hasta unos pocos nucleótidos (Yamada *et al.*, 2003; Mockler *et al.*, 2005; Yazaki *et al.*, 2007; Toledo-Arana *et al.*, 2009). Una limitación de los microarrays es que su diseño requiere conocer previamente el genoma, y además presentan un rango dinámico bastante limitado que se satura con facilidad (Royce *et al.*, 2007). A esto hay que sumarle que la comparación de la expresión entre experimentos requiere de complicados métodos de normalización (Yang *et al.*, 2002; Okoniewski

and Miller, 2006) y que el diseño y manufacturación de estos dispositivos tiene un elevado coste.

Por último, la técnica que ha revolucionado el campo de la transcriptómica en tan sólo un par de años, y que ha eclipsado a los tiling microarrays, es la secuenciación masiva de RNA (RNA-seq) (Wang *et al.*, 2009). La RNA-seq consiste en la secuenciación de cDNA o más recientemente en la secuenciación directa del RNA (Wang *et al.*, 2009; Thompson and Milos, 2011). Una vez obtenido el RNA total, se suele enriquecer la muestra retirando los rRNAs, ya que representan aproximadamente el ~80% del RNA total (Condon, 2007; Passalacqua *et al.*, 2009). Posteriormente se realiza una fragmentación por nebulización, y la ligación de los adaptadores en los extremos 5' y 3', lo que permite preservar la direccionalidad de la molécula de RNA (Wang *et al.*, 2009). Finalmente, el RNA se transforma en cDNA que luego es amplificado por PCR y secuenciado de forma masiva mediante técnicas como 454, Illumina o IonTorrent. Los avances en la preparación de la muestra, procesividad de la reacción de secuenciación, así como los algoritmos de concatenación, análisis y comparación de las lecturas están permitiendo evaluar los transcriptomas con una fiabilidad y profundidad inimaginable hace un par de años (Wang *et al.*, 2009). Gracias a ello, hoy se consigue determinar los extremos de todos los transcritos con una precisión de un nucleótido sin la necesidad de conocer la secuencia genómica (Bradford *et al.*, 2010). Además, variando la manera de realizar la librerías es posible determinar cómo los transcritos son procesados o modificados post-transcripcionalmente, e incluso determinar las estructuras

secundarias a escala global (Maher *et al.*, 2009; Roberts *et al.*, 2011). Gracias a ello, recientemente se ha puesto en evidencia nuevos mecanismos moleculares que muestran la importancia de la regulación post-transcripcional para el desarrollo de cualquier organismo.

1.2. LA REGULACIÓN POST-TRANSCRIPCIONAL

Ha pasado más de medio siglo desde que Francis Crick, en 1956, describiera el dogma central de la biología que explica el flujo de la información genética en un sistema biológico desde el DNA al RNA y luego a proteínas (Crick, 1970). Dentro de este esquema el RNA era considerado como un mero intermediario para el flujo de dicha información. Sin embargo, los estudios recientes (basados en tiling arrays y RNA-seq) sugieren que quizás se ha infravalorado el papel del RNA como elemento clave para controlar el momento adecuado en el que el código genético se transforma en proteínas. Dichos estudios plantean que los mecanismos de regulación post-transcripcional contribuyen tanto o más que la propia regulación transcripcional para determinar la concentración final de proteína (Vogel and Marcotte, 2012). Además, tienen la peculiaridad de ser más rápidos y económicos, desde el punto de vista metabólico, que la regulación mediada por proteínas. Así, la regulación post-transcripcional modula el nivel de traducción de una determinada proteína variando tanto la cantidad del mRNA disponible (mediante diversos procesos de procesamiento y degradación de RNAs) como la accesibilidad a la región

de unión al ribosoma. En los últimos años se ha demostrado que el papel del RNA como regulador es esencial en el control de numerosos procesos biológicos incluido la patogénesis, la resistencia a antibióticos, la supervivencia en el ambiente, la formación de biofilm entre otros (Mattick, 2005; Toledo-Arana *et al.*, 2007; Toledo-Arana *et al.*, 2009; Caldelari *et al.*, 2013). Los distintos mecanismos generales de regulación post-transcripcional descubiertos hasta ahora se detallan a continuación.

1.2.1. La degradación del RNA

En general, se tiende a pensar que la degradación del RNA cumple una mera función de limpieza de los mRNA que ya no son necesarios para la célula, reciclando los nucleótidos que lo componen. Sin embargo, recientemente se empieza a comprender que el procesamiento del RNA no es sólo la típica reacción enzima (RNase) vs sustrato (RNA). El procesamiento de RNA es un mecanismo complejo donde participan más factores de los inicialmente esperados. Estos factores adicionales pueden actuar como catalizadores o inhibidores de la reacción y por lo tanto determinan tanto la especificidad como la cinética de la reacción (Arraiano *et al.*, 2010). En general, los RNAs pueden ser estables o inestables en el tiempo (Deutscher, 2005). Los estables incluyen las especies estructurales, como los tRNAs y los rRNAs, que son los componentes esenciales para la maquinaria de traducción. También un RNA puede tener distintos grados de estabilidad en función de ciertas condiciones medioambientales. Estos

cambios en la estabilidad se dan mediante la participación de proteínas de unión a RNA, sRNAs reguladores y/o cambios en la estructura del propio RNA que modifican la acción de las ribonucleasas (Anderson and Dunman, 2009). La expresión y/o actividad de estos factores está influenciada por las señales ambientales, y en consecuencia el proceso de degradación y/o procesamiento de un determinado RNA cumple una función reguladora.

Otro argumento a favor del potencial regulador de esta maquinaria es que, a pesar de que las moléculas de RNA son químicamente similares en todos los organismos, las ribonucleasas y los factores accesorios han evolucionado de manera diferente. Aunque todos los organismos mantienen las ribonucleasas básicas (endo- y exonucleasas), su número y especificidad varía en función de las especies (Tabla 1). La Tabla 1 muestra las ribonucleasas más representativas presentes en bacterias modelo como son *Escherichia coli*, *Bacillus subtilis* y *Staphylococcus aureus* y explica brevemente sus funciones. En general, las endonucleasas suelen ser las encargadas de iniciar los procesos de degradación cortando las moléculas de RNA para liberar extremos que ahora son accesibles para las exonucleasas (Fig. 1). Pero además es importante tener en cuenta que un corte en un RNA no siempre implica degradación. Muchas moléculas de RNA necesitan ser procesadas para cumplir su función como por ejemplo los rRNAs. Este proceso también ocurre en ciertos mRNAs, que tienen que ser procesados de manera específica para generar un fragmento más estable y accesible a los ribosomas en contraposición de la molécula precursora que sería más inestable.

Tabla 1. Descripción de las ribonucleasas bacterianas

| RNasa | Bacteria | Descripción | Referencias |
|-------------------|---------------|---|--|
| Endoribonucleasas | RNasa III | ■ + ○ Corta RNA de cadena doble. Responsable de la degradación de los dúplex mRNA sRNA. Procesa 30S y el mRNA de <i>pnp</i> . Modulador de la respuesta de estrés. | Agrawal et al., 2003; Evguenieva-Hackenberg & Klug, 2000; Calin-Jageman & Nicholson, 2003 |
| | RNasa BN/Z | ■ + ○ Maduración de la 3' de tRNAs. | Nashimoto, 1997; Pellegrini et al., 2003 |
| | RNasa E | ■ + ○ Procesamiento y maduración de 5S, 16S, tRNA y RNase P. | Misra & Apirion, 1979; Li Z & Deutscher MP, 2002; Lin-Chao et al., 1999; Lundberg & Altman, 1995 |
| | RNasa J1 / J2 | ■ + ○ Función similar a RNasa E. También tienes función exonucleasa 5' - 3'. | Mathy et al., 2007; Bechthofer, 2009 |
| | RNasa G | ■ + ○ Segregación del cromosoma y división celular. Maduración 5' del 16S rRNA. | Okada et al., 1994; Li et al., 1999b; Wachi et al., 1999 |
| | RNasa H | ■ + ○ Degrada híbridos de RNA y RNA/DNA que se forman durante la replicación y reparación del DNA evitando especies aberrantes. | Itaya et al. 1999; Condon & Putzer, 2002; Worral & Luisi, 2007; Tadokoro & Kanaya, 2009 |
| | RNasa I | ■ + ○ Degrada RNA dejando mononucleótidos. Recicla ribonucleótidos del medio exterior. Defensa contra bacteriófagos. | Condon & Putzer, 2002; Zhu et al., 1990. |
| | RNasa P | ■ + ○ Ribozima considerada un vestigio del "mundo del RNA". Genera tRNAs y 4.5S rRNAs maduros. Procesa dúplex de mRNA, sRNAs, riboswiches y zonas intergénicas de operones policistrónicos. | Li & Altman, 2003; Altman, 2007; Liu & Altman, 2009 |
| Exoribonucleasas | Oligo-RNasa | ■ + ○ Degrada pequeñas cadenas RNA de 2 a 5 nt producidas por las otras RNAsas. Hidrólisis producida en sentido 3'-5'. | Ghosh & Deutscher, 1999 |
| | PNPasa | ■ + ○ Cataliza la degradación 3'-5' liberando nucleótidos difosfato hasta que encuentra RNA de cadena doble. Forma complejos de degradación en conjunción con otras RNAsas. | Spickler & Mackie, 2000; Liou et al., 2002; Lin & Lin-Chao, 2005 |
| | RNasa II | ■ + ○ Componente catalítico del exoma. Degrada progresivamente en dirección 3' - 5' hasta que encuentra RNA de cadena doble, liberando nucleótidos monofosfato. | Fraza, 2006; Spickler & Mackie, 2000; Zuo, 2006 |
| | RNasa R | ■ + ○ Degrada progresivamente en dirección 3'-5'. Puede digerir RNAs muy estructurados. Es capaz de digerir dúplex de RNA si tienen extremos 3' protuberantes. Tiene gran importancia en la degradación de RNAs poliadenilados. | Cheng & Deutscher, 2002; Vincent & Deutscher, 2006; Andrade et al., 2009a |

Referencia: ■ *E. coli* **+** *B. subtilis* **○** *S. aureus*

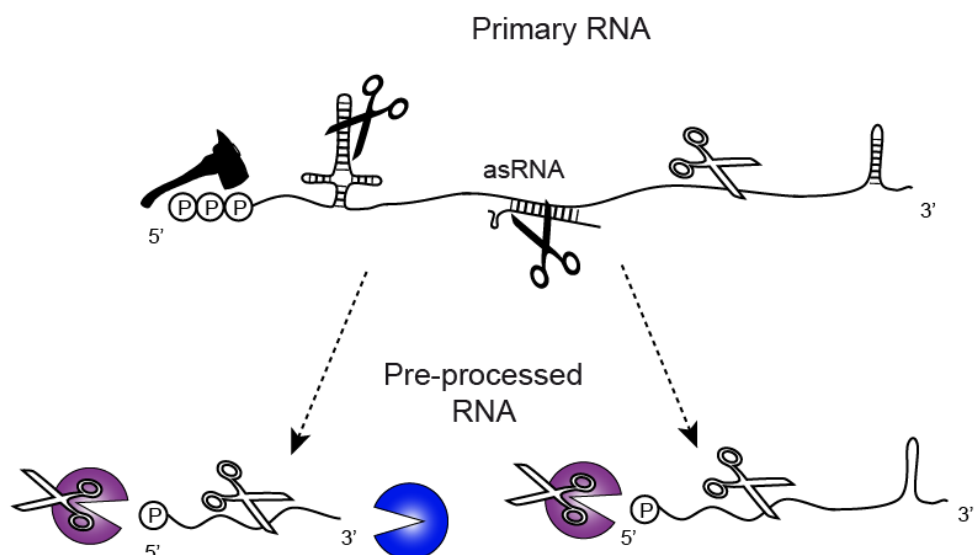


Figura 1. Representación esquemática que muestra de manera general el proceso de degradación de los mRNAs bacterianos. El proceso de degradación de un RNA primario comienza con digestiones endonucleolíticas que libera extremos 5' (monofosfato) y 3' que ahora son accesibles para las enzimas exonucleolíticas. Las enzimas endonucleolíticas pueden atacar regiones del RNA de cadena simple como las RNases E, G, J e Y (tijeras blancas) mientras que las regiones de cadena doble, que se pueden formar por estructuras internas o por la hibridación con RNA reguladores codificados en *cis* (RNA antisentidos) o en *trans* (RNA reguladores pequeños), son atacadas por RNase III (tijera negra). El extremo 5' del RNA primario contienen el trifosfato que es transformado a monofosfato por la enzima pirofosfotidrolasa (hacha). Una vez que el RNA primario ha sido preprocesado, las enzimas exonucleolíticas pueden atacarlo desde ambos extremos. La RNase J1 (morado) digiere el RNA desde el extremo 5' monofosfato mientras que PNPase digiere desde los extremos 3' que no estén estructurados. Si existiesen estructuras secundarias tipo stem-loops, la enzima PNPase necesita de la asistencia de proteínas RNA helicasas para continuar con el procesamiento del RNA.

1.2.2. Regulación mediada por RNAs

Los elementos reguladores contenidos en el RNA se pueden clasificar, de manera general, en elementos que actúan en “*trans*” o en “*cis*”. Los elementos *trans* son aquellos que están codificados en una región del genoma diferente a la del gen que regulan y comprenden normalmente los pequeños RNAs reguladores (sRNAs). En cambio, los elementos *cis* son estructuras o secuencias de RNA que se encuentran localizadas en la misma región del gen que regulan. Dentro de esta categoría se incluyen elementos reguladores como los RNA antisentidos y las regiones no traducidas del RNA mensajero localizadas tanto en el extremo 5' (5'-UTR) como el 3' (3'-UTR). A continuación se describen los elementos de RNA reguladores más comunes.

1.2.3. RNAs reguladores pequeños (sRNAs)

Los sRNAs son un grupo heterogéneo de moléculas que regulan rápidamente la fisiología de la bacteria atendiendo a todo tipo de señales (Gottesman and Storz, 2011). Los sRNAs son RNAs cortos que generalmente no codifican para ninguna proteína. Estos sRNAs pueden ser generados desde un promotor propio, o de un mRNA mediante un procesamiento diferencial o una parada de la transcripción dependiente de un metabolito como en los riboswitches (Novick *et al.*, 1993; Loh *et al.*, 2009; Gottesman and Storz, 2011; Chao *et al.*, 2012). Los sRNAs han sido implicados en diversas funciones biológicas que incluyen la regulación del

metabolismo (catabolismo y anabolismo), inhibición del crecimiento mediante el control de toxinas, virulencia, formación de biofilm, adaptación al stress y/o condiciones ambientales como la temperatura, la densidad celular, etc. De manera general, la gran mayoría de los sRNAs actúan apareándose con el mRNA diana para modificar su traducción modulando la acción de los ribosomas y/o las ribonucleasas. Existen también sRNAs con capacidad para unirse a ciertas proteínas reguladores como CsrA con el fin de modificar su actividad (Massé *et al.*, 2007; Toledo-Arana *et al.*, 2007; Desnoyers *et al.*, 2009; Romby and Charpentier, 2009; Jane M Liu and Camilli, 2010; Gottesman and Storz, 2011; Sharma and Heidrich, 2012; Caldelari *et al.*, 2013).

1.2.4. RNAs antisentidos

Dentro de los elementos reguladores en *cis*, hoy sabemos que los RNA antisentidos (asRNAs) son los más numerosos. Las nuevas técnicas transcriptómicas han revelado que los asRNAs cubren gran parte del genoma de diversos organismos (Lior David *et al.*, 2006; Kapranov *et al.*, 2007; Henz *et al.*, 2007; Y He *et al.*, 2008; Lasa *et al.*, 2011; Sesto *et al.*, 2013; Lybecker *et al.*, 2014). Como los RNAs antisentidos se transcriben desde la cadena opuesta al gen que regulan, tienen una complementariedad total con el mRNA producido desde la cadena sentido. Es importante entender que la región de solapamiento se puede generar mediante diversas configuraciones cromosómicas. Es decir, además de los

RNA antisentidos que se generan desde un promotor propio, también es posible crear zonas de solapamiento con regiones 3'-UTRs provenientes de genes convergentes que carecen de terminador de transcripción o regiones 5'-UTRs de genes divergentes, en donde el promotor de uno está codificado dentro (o por detrás) de la secuencia del otro (Fig. 2). En otros casos particulares, es posible que el mRNA de un gen codificado en medio de un operón, pero en sentido contrario al resto de genes, solape con el mRNA de dicho operón (Fig. 2) (Georg and Hess, 2011; Lasa *et al.*, 2011; Lasa *et al.*, 2012; Lybecker *et al.*, 2014). Aunque la función de la mayoría de estos asRNAs es desconocida, se especula que pueden actuar interfiriendo en el proceso de la transcripción, alterando la estabilidad del mRNA y/o inhibiendo la traducción (Thomason and Storz, 2010; Georg and Hess, 2011).

Estudios recientes realizados en nuestro laboratorio han demostrado que *Staphylococcus aureus* codifica cientos de estos RNAs reguladores que cubren gran parte del genoma (Lasa *et al.*, 2011). Basado en la secuenciación masiva de las fracciones de RNAs pequeños (menores a 50 nt) y RNA largos, detectamos la presencia de una colección de fragmentos de RNA (largo promedio de 20 nt) que se distribuía simétricamente en ambas cadenas del genoma. Llamativamente, la localización genómica de estos RNA cortos se correlacionaba con la zona de regiones de solapamiento entre los RNA largos.

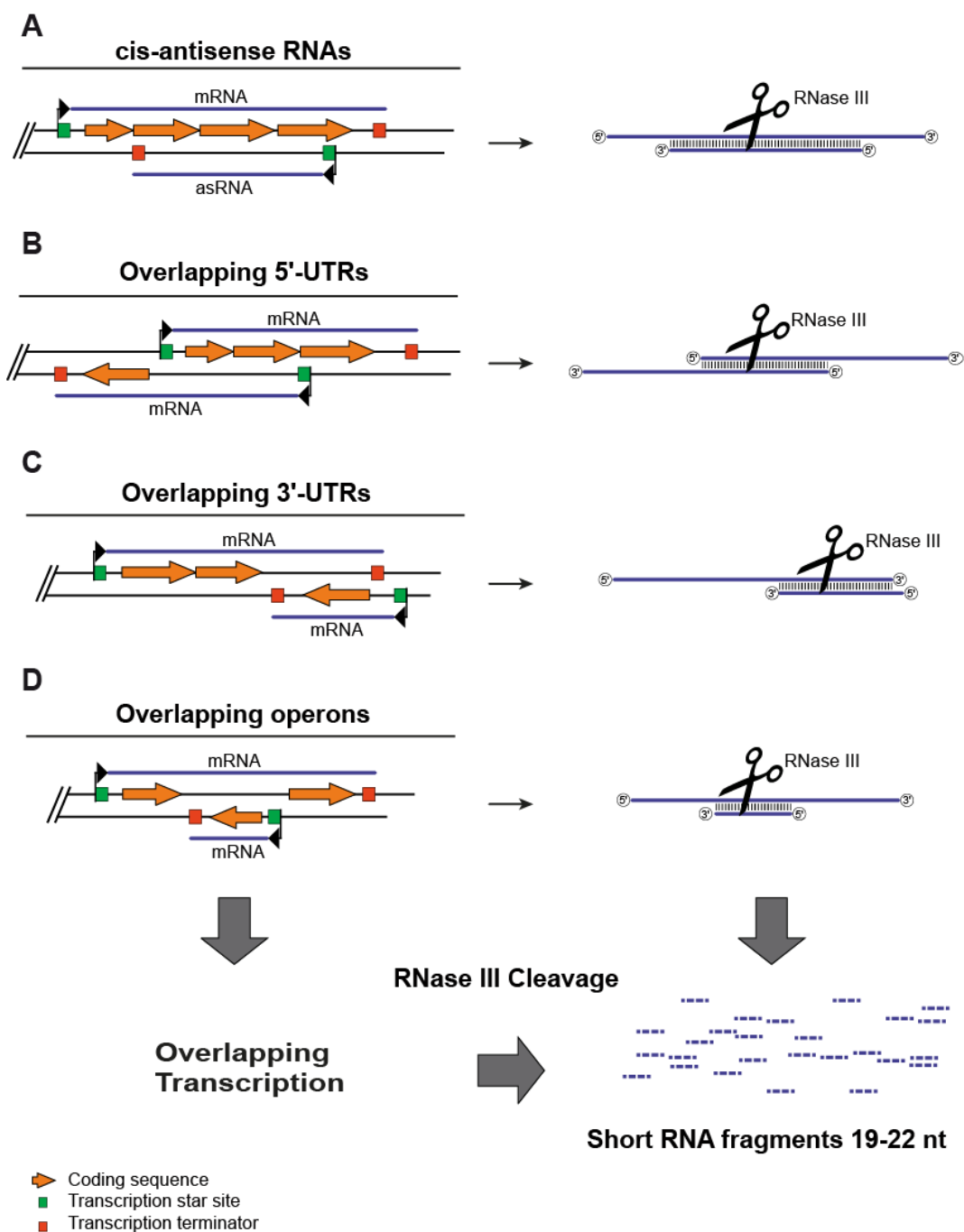


Figura 2. Representación esquemática de los distintos mecanismos para generar RNAs solapantes. Los RNA solapantes se pueden generar mediante distintas configuraciones cromosómicas: **(A)** los RNA antisentidos (asRNA) se generan desde un promotor propio codificado en la cadena contraria a una ORF y pueden afectar a uno o varios genes. **(B)** Las regiones 5' solapantes se generan cuando el promotor de un gen se encuentra codificado dentro (o detrás) de la ORF situada en la cadena opuesta del DNA. **(C)** Las regiones 3' solapantes ocurren cuando el terminator de transcripción está codificado dentro (o por detrás) del gen situado en la cadena opuesta del DNA. **(D)** El mRNA de un gen codificado en medio de un operón, pero en sentido contrario al resto de genes, genera un solapamiento con el mRNA de dicho operón. Sin importar la manera que se generan las regiones de RNA antisentido, todas son procesadas por RNase III, produciendo como resultado fragmentos de RNA de unos 19 a 22 nt (Lasa *et al.*, 2011).

La ausencia de estos fragmentos cortos de RNA en la secuenciación de la fracción de RNAs pequeños en el mutante en RNase III, demostró que dichos RNAs cortos eran el resultado del procesamiento llevado a cabo por esta enzima que reconocía la región de RNA cadena que se generaba por el solapamiento de los RNAs sentido y antisentido (Lasa *et al.*, 2011). Este proceso también ocurría en diferentes bacterias Gram-positivas como *Listeria monocytogenes*, *Bacillus subtilis* y *Enterococcus faecalis*. Posteriormente, de manera similar a nuestros resultados, dos grupos han demostrado la participación de RNase III en el procesamiento global de los asRNAs en *S. aureus* y *Escherichia coli* (Lioliou, Sharma, *et al.*, 2012; Lybecker *et al.*, 2014). En el primer caso, mediante secuenciación masiva, se identificaron todos los RNA asociados a una proteína RNase III modificada, la cual permitía unirse a los RNAs de cadena doble (dsRNAs) pero no digerirlos. En el segundo caso, Lybecker *et al* (Lybecker *et al.*, 2014) inmunoprecipitaron “*in vivo*” los asRNAs ayudados por un anticuerpo que se une a los dsRNAs, para posteriormente secuenciar el RNA obtenido. Estos autores observaron que era más sencillo recuperar los dsRNAs en ausencia de RNase III (Lybecker *et al.*, 2014). En su conjunto, estos tres trabajos demostraban que tanto el RNA sentido como el RNA antisentido pueden coexistir en una misma célula y apoyan la teoría del excludón como mecanismo regulador de la expresión génica (Sesto *et al.*, 2013). Esta teoría propone que genes adyacentes (codificados uno opuesto al otro) y generalmente de funciones biológicas opuestas regulan su expresión mediante la formación de RNAs antisentidos, que serían procesados por

RNase III. En consecuencia, la expresión de uno excluye la expresión del otro (y viceversa) y sólo cuando el mRNA se exprese en mayor medida que su antisentido podrá ser traducido a proteína (Sesto *et al.*, 2013).

1.2.5. Regiones no traducidas (UTRs) del mRNA

Con la evolución del conocimiento de los procesos de regulación post-transcripcional hemos comprendido que los mRNAs, además de ser los intermediarios en la transferencia de la información genética necesarios para producir las proteínas, pueden desempeñar múltiples funciones reguladoras. Una molécula de mRNA bacteriano está formada básicamente por una región codificante (que contiene los codones para la traducción a proteínas) y dos regiones no codificantes, una localizada en el extremo 5' y la otra en el 3' del mRNA. La 5'-UTR comprende desde el inicio de transcripción determinado por la secuencia promotora hasta el primer codón de inicio de la traducción. En cambio, la 3'-UTR comprende desde el último codón de traducción hasta el último nucleótido del mRNA determinado por la terminación de la transcripción. El proceso de terminación puede depender de la proteína Rho o ser independiente de ésta (Peters *et al.*, 2011; Jiankui He *et al.*, 2012; Régnier and Hajnsdorf, 2013; Boudvillain *et al.*, 2013). Si es independiente, la terminación se genera por la formación de una estructura secundaria en el RNA de tipo “*stem-loop*” que fuerza la liberación de la RNA polimerasa (Wilson and Hippel, 1995).

Hasta el momento, el estudio de la regulación post-transcripcional mediada por las regiones no traducidas del mRNA ha estado centrado principalmente en los extremos 5', probablemente porque estas regiones contienen el sitio de unión al ribosoma (RBS). En ésta región se han descrito diversos elementos reguladores como los riboswitches y termosensores. Además, las 5'-UTRs son el blanco de acción de numerosos sRNAs que se aparean en esta zona como así también de proteínas de unión a RNA como CsrA. En todos los casos, la función de estos elementos es modular la síntesis de la proteína modificando la accesibilidad del ribosoma a la RBS y/o la estabilidad del mRNA (Waters and Storz, 2009).

Los riboswitches consisten fundamentalmente en dos dominios: el adaptámero o domino sensor que es reconocido por un ligando concreto y la plataforma de expresión o dominio efector responsable del cambio de expresión de la proteína (Breaker, 2012; Serganov and Nudler, 2013). Cuando el dominio sensor se une al ligando correspondiente, provoca un cambio en la estructura secundaria del dominio efector, haciendo que la transcripción o la traducción se vean afectadas. Según el caso, una parada prematura en la transcripción genera sRNAs. Se ha demostrado que estos sRNAs derivados de los riboswitches también pueden actuar en *trans* regulando otros genes (Loh *et al.*, 2009). En general, los riboswitches se encuentran por delante de genes que codifican proteínas necesarias para el metabolismo y los ligandos suelen ser compuestos relacionados con la vía metabólica en la que participa la proteína cuya expresión depende del

riboswitch (Breaker, 2012). En la Figura 3A, se muestran ejemplos de distintos mecanismos utilizados por los riboswitches para controlar la expresión proteica.

Otro mecanismo interesante codificado en las 5'-UTRs son las ribozimas (Fig. 3B). En este caso, la unión del metabolito a dicha región provoca un cambio conformacional en la estructura secundaria que induce un autoprocésamiento del mRNA generando mRNAs alternativos o promoviendo su degradación (Klein and Ferré-D'Amaré, 2006; Lee and Groisman, 2010).

Además de estos mecanismos, algunas 5'-UTRs forman estructuras secundarias que se ven alteradas por variaciones de temperatura (Fig. 3C). Así, estas estructuras pueden actuar como termómetros o termosensores que sólo dejan expresar la proteína codificada por el mRNA a una temperatura en particular sin la necesidad de unir un ligando específico (Narberhaus, 2010; Kortmann and Narberhaus, 2012). Por ejemplo, la región 5'-UTR de *prfA*, el mRNA que codifica para el regulador de virulencia de *Listeria monocytogenes*, contiene un *stem-loop* que bloquea el acceso al ribosoma a temperatura ambiente. En cambio, cuando la bacteria se encuentra en el hospedador, a 37°C, el *stem-loop* se disocia permitiendo la expresión del regulador PrfA, que activa toda la batería de proteínas necesarias para la infección (Johansson *et al.*, 2002).

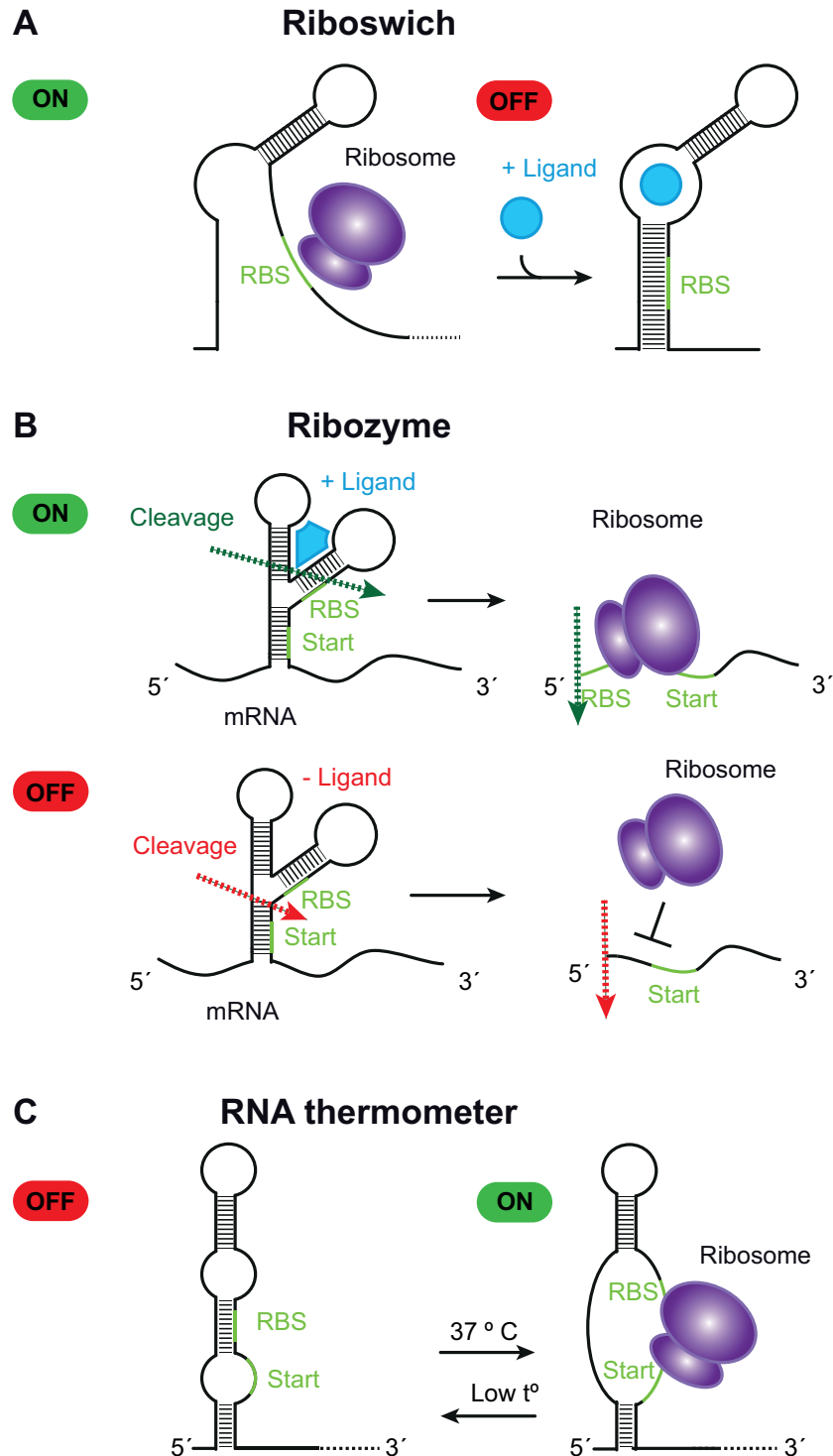


Figura 3. Mecanismos de regulación post-transcripcionales situados en 5'-UTRs. (A)

Los *riboswitches* modulan la traducción de las proteínas codificadas en su mRNA mediante la atenuación de la transcripción o por el bloqueo del acceso del ribosoma mediante un cambio estructural de la 5'-UTR debido a la presencia/ausencia de un determinado metabolito. **(B)** Las ribozimas, tras la unión de su ligando, cambia la estructura secundaria de la 5'-UTR y en consecuencia cambia el sitio de procesamiento generando un mRNA que es ahora reconocido por el ribosoma. **(C)** Los termosensores son RNAs que modulan la estructura de la región en función de la temperatura. Los ejemplos clásicos son aquellos que, a baja temperatura, la estructura oculta la RBS mientras que a una temperatura más alta, la estructura se desorganiza para liberar la RBS y facilitar la traducción.

Con respecto a las regiones 3'-UTRs, hasta hace poco no habían sido consideradas como elementos con potencial para regular la expresión de genes en bacterias. En cambio, en eucariotas, estas regiones son esenciales para la correcta expresión de proteínas. Las 3'-UTRs eucariotas contienen secuencias específicas para reclutar diversos factores reguladores. Por ejemplo, la cola de poli(A) sirve como sitio de reconocimiento para diversas proteínas de unión a RNA, algunas necesarias para que los mRNAs se exporten fuera del núcleo, y otras para promover la traducción o aumentar la estabilidad de los mRNAs (Mignone *et al.*, 2002; Matoulkova *et al.*, 2012). Otras secuencias contenidas en las 3'-UTR eucariotas son los elementos AU-rich y GU-rich (AREs y GREs respectivamente) que son reconocidos por proteínas que promueven la degradación del mRNA en respuesta a señales intracelulares y extracelulares muy específicas (Vlasova *et al.*, 2008; Khabar, 2010; Halees *et al.*, 2011). Además las 3'-UTRs son blanco para la acción de los microRNAs, que producen el silenciamiento de genes mediante la desestabilización del mRNA (Brodersen and Voinnet, 2009). El descubrimiento de mutaciones en las regiones 3'-UTRs, que causan un cambio en la estructura secundaria o la desestabilización de los transcritos, relacionadas con ciertas enfermedades graves como Alzheimer, esclerosis múltiple, diabetes y varios tipos de cánceres, pone en evidencia la importancia de la regulación mediada por estas 3'-UTRs (Jian-Min Chen *et al.*, 2006; Sandberg *et al.*, 2008; Mayr and Bartel, 2009). También se ha especulado que la evolución de estos mecanismos de regulación está

relacionado con la evolución de los organismos en el árbol de la vida. Dicho de otra manera, el largo de las 3'-UTRs y por lo tanto su capacidad para albergar elementos reguladores tiene relación con la especiación. Así la longitud de las 3'-UTRs aumentan con la complejidad de los organismos, siendo la especie humana la que contiene las de mayor tamaño (Mazumder *et al.*, 2003; C Y Chen *et al.*, 2012).

Por el contrario, se pensaba que en organismos simples como los procariotas, con 3'-UTR mucho más cortas, la función de estas regiones no iba más allá de contener la señales de terminación de la transcripción y prevenir la acción de los ribonucleasas. Sin embargo, los recientes mapas transcriptómicos realizados en algunas bacterias han revelado ciertas evidencias que nos han llevado a plantear que quizás las 3'-UTRs pueden cumplir funciones reguladores similares a las eucariotas. Por ejemplo, en *Bacillus subtilis* se identificaron nueve genes con 3'-UTR largas (~220 nt) muy conservadas. Llamativamente, estos mRNAs codifican para proteínas asociadas con la membrana plasmática lo que sugiere una regulación post-transcripcional común para estos genes (Rasmussen *et al.*, 2009). En *L. monocytogenes*, *S. aureus* y *Mycobacterium tuberculosis* se han encontrado cientos de genes convergentes en donde sus 3'-UTRs solapan formando antisentidos que son procesados mediante RNase III (Toledo-Arana *et al.*, 2009; Arnvig *et al.*, 2011; Lasa *et al.*, 2011). La determinación global del largo de las 3'-UTRs tanto en *S. aureus* como en *Porphyromonas gingivalis* reveló que, en general, las 3'-UTRs son más largas que las 5'-UTRs, confirmando la idea de que estas regiones pueden contener más

información de la esperada (Broeke-Smits *et al.*, 2010; Lasa *et al.*, 2011; Høvik *et al.*, 2012). De hecho, la secuenciación masiva los RNA coimmunoprecipitados con Hfq de *Salmonella typhimurium*, una chaperona de RNA que participa en la regulación mediada por sRNAs, reveló que esta proteína se une a varias 3'-UTRs (Sittka *et al.*, 2008). En relación con estos resultados, se descubrió que las regiones 3'-UTRs pueden generar sRNAs, ya sea mediante procesamiento del mRNA por alguna RNasa o porque existe un promotor alternativo en la zona de la 3'-UTR. Estos sRNAs pueden actuar en *trans* sobre determinados mRNAs para regular su expresión (Kawano, 2005; Chao *et al.*, 2012). Como resultado, estos mRNAs son considerados moléculas multifuncionales porque tienen la capacidad para producir una proteína, pero al mismo tiempo contienen elementos reguladores que controlan la expresión de otras proteínas.

El paradigma de la multifuncionalidad de un mRNA bacteriano se representa con el RNA regulador RNAIII de *S. aureus*. RNAIII es uno de los RNA reguladores más estudiados y fue el primero descrito en esta bacteria (Novick *et al.*, 1993). El RNAIII es la molécula efectora del sistema de *quórum sensing* Agr. El operón *agr* contiene cuatro genes que se encargan de sensar y responder ante la densidad celular (Fig. 4). El gen *agrD* produce el péptido de autoinducción que es continuamente secretado al medioambiente. Cuando llega a una concentración crítica el sistema de dos componentes AgrA-AgrC se activa desencadenando la expresión de RNAIII (Novick, 2003). RNAIII es un mRNA mensajero que codifica para la proteína δ -hemolisina.

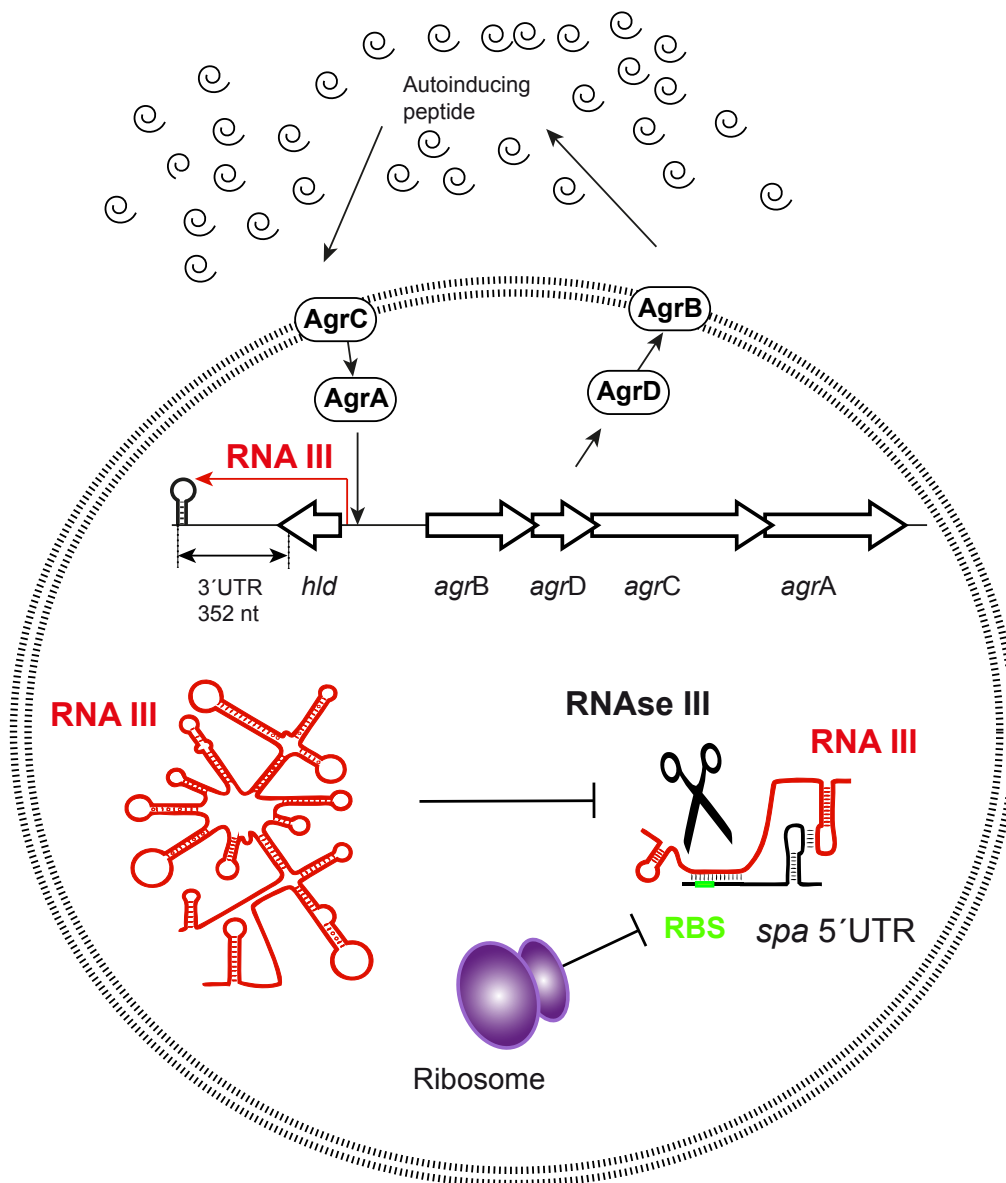


Figura 4. Esquema del sistema de *quorum sensing* *agr*. El sistema *agr* está formado por un operón de cuatro genes (*agrADBC*) y una molécula de RNA efectora (RNAIII). AgrA y AgrC son el sistema de dos componentes que detecta la presencia del péptido inductor que es procesado a partir de AgrD por la proteína de membrana AgrB. Cuando el péptido es liberado al exterior celular, el sistema de dos componentes se activa para transcribir RNAIII. Este RNA codifica la δ -hemolisina (*hld*) y además contiene una 3'-UTR larga de 352 nt que forma una compleja estructura secundaria. RNAIII es capaz de unirse a diversas 5'-UTRs gracias a los motivos UCCCC que hibridan con las RBSs. Esto bloquea el acceso del ribosoma y forma una región de RNA de cadena doble en la región que es procesada por RNasa III (Romby *et al.*, 2009; Felden *et al.*, 2011).

Además, este mRNA presenta una 3'-UTR larga de 352 nt mediante la cual regula de manera post-transcripcional una serie de genes implicados en la virulencia y la producción de toxinas (Huntzinger *et al.*, 2005; Boisset *et al.*, 2007; Romby and Charpentier, 2009; Chevalier *et al.*, 2010; Felden *et al.*, 2011). El mecanismo consiste en el emparejamiento de RNAIII, a través de diferentes *stem-loops* localizados en el mRNA, con la 5'-UTR de los mRNAs dianas. Esta interacción en *trans* entre los mRNAs produce la oclusión de la zona de unión del ribosoma y promueve el procesamiento mediado por RNase III de los mRNAs para inhibir la traducción de la proteína.

Todo estas evidencias demuestran que los mecanismos de regulación mediados por 3'-UTRs, ahora visto como excepcionales, en realidad pueden ser más generales de lo previsto. La utilización de las regiones 3'-UTRs como elementos reguladores puede tener ciertas ventajas, por ejemplo, tanto la secuencia como la estructura secundaria del mRNA en esta región no esta condicionada por el proceso de traducción. Por lo tanto esta región puede evolucionar de manera diferente a la región codificante permitiendo generar funciones adicionales para ese mRNA.

1.3. STAPHYLOCOCCUS AUREUS COMO MODELO DE ESTUDIO

S. aureus es una de las bacterias modelo para el estudio de los mecanismos de regulación mediados por RNA. Como se ha mencionado anteriormente, en esta bacteria se han descrito numerosos mecanismos

que involucran RNAs reguladores, tanto pequeños RNAs como RNAs antisentidos, y que han contribuido significativamente al avance en el conocimiento de los procesos de regulación post-transcripcional bacteriana (Toledo-Arana *et al.*, 2007; Desnoyers *et al.*, 2009; Romby and Charpentier, 2009; Jane M Liu and Camilli, 2010; Gottesman and Storz, 2011; Lasa *et al.*, 2011; Lioliou, Sharma, *et al.*, 2012; Caldelari *et al.*, 2013). Además debido a su enorme relevancia en el ámbito de la sanidad humana y animal hace que esta bacteria sea un modelo de estudio inmejorable.

1.3.1. Características de *S. aureus*

S. aureus fue descrito por primera vez por Ogston (Ogston, 1881) en el pus que drenaba de un absceso infectado. El nombre de esta especie proviene del binomio griego *staphylé* y *coccus*, debido a que tienen aspecto de cocos que forman una especie de racimo de uvas bajo la luz del microscopio, mientras que el característico color dorado de las colonias que forma le dan el nombre de *aureus*. *S. aureus* es una bacteria Gram-positiva, anaerobia facultativa ubicua que se encuentra ampliamente distribuida por todo el mundo colonizando tanto humanos como animales (Lowy, 1998; Leonard and Markey, 2008; Gordon and Lowy, 2008; Salyers and Whitt, 2011). Aproximadamente el 20% de los adultos son portadores persistentes en sus fosas nasales y el 30% son colonizados ocasionalmente (Aly *et al.*, 1980; Wertheim *et al.*, 2005; Mulcahy *et al.*, 2012; Yan *et al.*, 2013). Además puede colonizar otras zonas como la piel, oído, axilas, ingle y tracto intestinal. A partir de estos reservorios naturales donde persiste, en

determinadas circunstancias, puede atravesar la barrera epitelial y causar diversas infecciones. Estas infecciones van desde las más leves como los accesos cutáneos, impétigo, foliculitis y mastitis; hasta las más graves como las bacteremias, meningitis, osteomielitis, neumonía, endocarditis, que pueden causar la muerte del hospedador (Lowy, 1998; Gordon and Lowy, 2008).

1.3.2. Patogenia de *S. aureus* y multi-resistencia antibiótica

La patogenidad de esta bacteria depende de diversos factores de virulencia como son las proteínas de superficie que ayudan a la adhesión, las enterotoxinas, las proteasas y los superantígenos (Foster and Hook, 1998). Asimismo, puede secretar una serie de proteínas que inhiben la activación del complemento y la quimiotaxis de los neutrófilos evitando la respuesta del sistema inmune (Foster, 2005; Hanke and Kielian, 2012; Valle *et al.*, 2013). Además, mediante la acción de Proteína A (una MSCRAMM, componente de la superficie microbiana que reconocen moléculas adhesivas), que une a los anticuerpos IgG para contrarrestar la opsonización, genera un efecto antifagocítico que evita que los macrófagos puedan detectar eficientemente a las bacterias durante el proceso de infección (Hall *et al.*, 2003).

Algunas estirpes de *S. aureus* pueden llevar factores de virulencia alternativos que normalmente son adquiridos luego de la infección por algunos bacteriófagos. Por ejemplo, las proteínas que promueven la

colonización de los tejidos del huésped como la Panton-Valentine leukocidine (Labandeira-Rey *et al.*, 2007), exotoxinas que dañan los tejidos del huésped para convertirlos en nutrientes asimilables por la bacteria (Dinges *et al.*, 2000), enterotoxinas que causan contaminaciones alimentarias y producen el síndrome de shock tóxico (Pinchuk *et al.*, 2010).

La virulencia de esta bacteria se demuestra por el hecho de que, antes de la introducción de la penicilina, el 80% de los individuos infectados por *S. aureus* morían a causa de la infección. Lamentablemente, con la misma facilidad que *S. aureus* puede adquirir factores de virulencia, puede integrar en su genoma genes de resistencia a diversos antimicrobianos. En los últimos años, ha adquirido resistencias a casi la totalidad de los antibióticos que normalmente se utilizan en clínica. En 1945, tan solo cuatro años después de la introducción de la penicilina, se aisló el primer *S. aureus* resistente a la penicilina, que la hidrolizaba por la acción de una β -lactamasa (Spink, 1954). En 1961 (Jevons *et al.*, 1961). Dos años después de la introducción de la meticilina en los hospitales, *S. aureus* desarrolló resistencia a este fármaco por la adquisición del gen *mecA*. El gen *mecA* codifica una proteína fijadora de penicilina llamada PBP2a que presenta una baja afinidad por β -lactámicos (Chambers *et al.*, 1985). La adquisición de esta resistencia ha servido para acuñar las siglas SARM por *S. aureus* resistentes a meticilina (o MRSA en sus siglas en inglés). Hoy los MRSA son un grupo de estirpes que incluyen resistencias a meticilina, nafcilina y oxacilina que normalmente están reguladas por una serie de genes que se encuentran en la región del cromosoma de *S. aureus* llamada casete

cromosómico estafilocócico (SSC). El SCC es un elemento genético móvil de transferencia horizontal intra e inter especies (Hanssen and Ericson Sollid, 2006; Chambers and DeLeo, 2009).

Adquisiciones sucesivas de mecanismos de resistencia a todas las clases de agentes antibióticos utilizados tanto en salud humana como en sanidad animal ha convertido el tratamiento y el control de las infecciones por *S. aureus* en un enorme problema sanitario. Además, la reciente emergencia de *S. aureus* resistentes a vancomicina, el tratamiento estándar para los MRSA, nos deja sin ningún antibiótico para tratar estas cepas multiresistentes (Jones, 2008). Todos estos factores hacen que *S. aureus* sea una bacteria extremadamente versátil y por ello sea considerada como uno de la patógenos hospitalarios más temidos a nivel mundial.

1.3.3. La capacidad para formar biofilms de *S. aureus*

Otra factor añadido que aumenta considerablemente el éxito de *S. aureus* como patógeno es su capacidad para formar biofilms. Aproximadamente el 75% de las infecciones de implantes médicos están causadas por alguna especie de estafilococos (Montanaro *et al.*, 2011). Este tipo de infecciones son difíciles de erradicar mediante los tratamientos antibióticos tradicionales, debido a que la bacteria en el interior del biofilm es más resistente a los antimicrobianos y a la acción del sistema inmune. Por ello, en la mayoría de las ocasiones es necesario sustituir el dispositivo médico para eliminar la infección (Fitzpatrick *et al.*, 2005). Esto aumenta

considerablemente tanto el riesgo para el paciente como los costes económicos de las hospitalizaciones.

Los biofilms por *S. aureus* también son un problema grave en producción animal. La capacidad de estas bacterias para formar biofilms sobre tejidos vivos como el epitelio de la glándula mamaria juega un papel fundamental en el desarrollo de infecciones intramamarias, que con frecuencia acaban desarrollando mastitis crónicas (Baselga *et al.*, 1993; Cucarella *et al.*, 2001). Del mismo modo que en los hospitales, estas infecciones son muy difíciles de eliminar y, en consecuencia, los animales tienen que ser sacrificados con la consiguiente pérdida económica para el ganadero (Barkema *et al.*, 2006; Nickerson, 2009).

1.3.3.1. El proceso de formación de biofilms en *S. aureus*

El biofilm (o biopelícula) se define como una comunidad de microorganismos adherida a una superficie inerte o tejido vivo, que produce una matriz extracelular que los mantiene unidos (Costerton *et al.*, 1978; Flemming and Wingender, 2010). El proceso de formación del biofilm es un ciclo compuesto de varias fases bien diferenciadas (Götz, 2002; O'Gara, 2007; Boles and Horswill, 2011). La primera es la fase de adhesión, en donde las bacterias se adhieren a una determinada superficie con la ayuda de proteínas que pueden estar unidas covalentemente o no a la superficie bacteriana (por ejemplo las MSCRAMMs). Estos factores adhesivos permiten que *S. aureus* pueda unirse a fibronectina, fibrinógeno, elastina y

colágeno promoviendo la adhesión a las células mucosas y a los tejidos (Foster and Hook, 1998). Con la misma eficiencia, la bacteria también puede unirse a cualquier implante médico que normalmente están recubiertos por dichas proteínas del hospedador (Montanaro *et al.*, 2011). Durante la siguiente fase, la de acumulación, las bacterias empiezan a dividirse activamente y a sintetizar la matriz del biofilm. Posteriormente, en la fase de maduración, se crea una estructura tridimensional con canales intersticiales por los que circula el agua y los nutrientes (Costerton *et al.*, 1999; Stoodley *et al.*, 2002). El ciclo se cierra cuando las bacterias se liberan, a partir de esta estructura, por factores mecánicos y/o moleculares, permitiendo la colonización de nuevas superficies (Boles and Horswill, 2011).

1.3.3.2. La matriz del biofilm en *S. aureus*

En la mayoría de los biofilms, los microorganismos representan el 10 % de la materia seca mientras que el 90 % restante es la matriz del biofilm (Flemming and Wingender, 2010). La matriz protege a los microorganismos de la desecación, agentes oxidantes o reductores, cationes metálicos, radiación ultravioleta, de otros microorganismos depredadores y de la acción del sistema inmune. Permite la retención de agua y enzimas extracelulares convirtiendo el biofilm en un sistema de digestión externo donde las bacterias pueden absorber directamente los nutrientes necesarios. La matriz del biofilm puede estar compuesta principalmente por

polisacáridos extracelulares (EPS), proteínas, ácidos nucleicos y/o lípidos (Cucarella *et al.*, 2001; Götz, 2002; Izano *et al.*, 2008; Vergara-Irigaray *et al.*, 2008; Flemming and Wingender, 2010).

En el caso de los biofilms de *S. aureus*, se ha descrito que esta bacteria es capaz de sintetizar principalmente dos tipos de matriz, una formada por proteínas u otra por polisacáridos. De hecho, una misma cepa puede alternar entre sintetizar una matriz u otra en función de las condiciones ambientales (Toledo-Arana *et al.*, 2005; N Merino *et al.*, 2009; Vergara-Irigaray *et al.*, 2009). La matriz proteica puede ser ensamblada mediante proteínas de superficie como Bap, FnBPs, SasG o Proteína A (Cucarella *et al.*, 2001; Corrigan *et al.*, 2007; Kuroda *et al.*, 2007; Shanks *et al.*, 2008; O'Neill *et al.*, 2008; N Merino *et al.*, 2009; Vergara-Irigaray *et al.*, 2009), mientras que la matriz polisacáridica esta formada principalmente por el exopolisacárido PIA-PNAG (Mack *et al.*, 1996; McKenney *et al.*, 1998). Este polímero se forma por la unión lineal de numerosos residuos de β -1,6 N-acetil-glucosamina, parcialmente deacetilados y por tanto cargados positivamente. Esta carga contribuye a la adhesión de la matriz a la membrana bacteriana y otras superficies cargadas negativamente. El PIA/PNAG es sintetizado y secretado gracias a la acción de cuatro proteínas codificadas por el operón *icaADBC* (Fig. 5A y 5B) (Heilmann *et al.*, 1996).

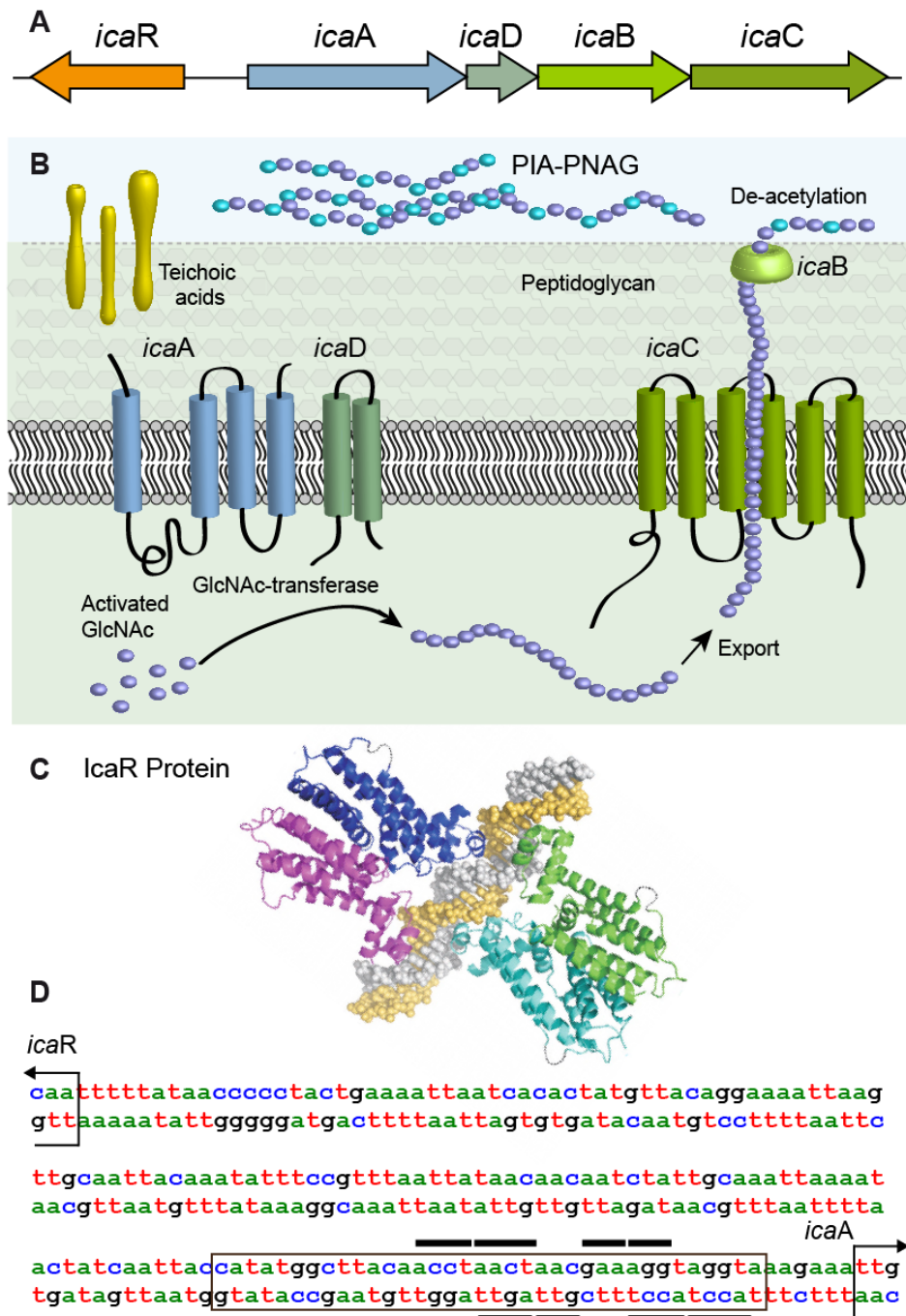


Figura 5. Representación esquemática de los componentes que participan en la síntesis del exopolisacárido PIA-PNAG de *S. aureus*. (A) Organización del operon *icaADBC* y del gen *icaR*. (B) Localización y función de las proteínas implicadas en la síntesis del PIA-PNAG (Heilmann *et al.*, 1996; Otto, 2009). (C) Modelo de dos dímeros de proteína IcaR unidos a la región promotora de *icaA* en base a la estructura cristalizada (PDB: 2ZCN) (Jeng *et al.*, 2008). (D) Región intergénica entre el operon *icaADBC* e *icaR* de *Staphylococcus aureus*. La flechas señalan el inicio de las proteínas IcaR e IcaA. El rectángulo señala la zona de unión de IcaR descrita por Jefferson y col., mientras que las líneas gruesas, las descrita por Jeng y col. (Jefferson *et al.*, 2003; Jeng *et al.*, 2008).

Las proteínas IcaA, IcaD e IcaC se encuentran embebidas en la membrana bacteriana mientras IcaB se encuentra en el exterior (Gerke, 1998). IcaA e IcaD trabajan conjuntamente como N-acetilglucosamintransferasas para sintetizar oligómeros de N-acetil-glucosamina de menos de 20 residuos. La proteína de membrana IcaC se encarga de unir los oligómeros en cadenas mas largas y transportarlos al exterior (Gerke, 1998). Finalmente la proteína IcaB es responsable de la deacetilación de un 15-20% de los residuos de PIA/PNAG (Vuong *et al.*, 2004).

1.3.3.3. La regulación de la síntesis del exopolisacárido PIA-PNAG

El principal regulador que controla la expresión de las proteínas codificadas en el operón *icaADBC* es la proteína IcaR, que actúa como un regulador transcripcional negativo. El gen *icaR* se encuentra localizado en el mismo locus pero en sentido contrario a dicho operón (Fig. 5A). IcaR es una proteína de unión a DNA que forma homodímeros que se unen a una región localizada a 42 nt del gen *icaA* e impide la transcripción del operón (Fig. 5C y 5D) (Jefferson *et al.*, 2003). IcaR pertenece a la familia de los represores de tetraciclina TetR (Jeng *et al.*, 2008). Este tipo de reguladores cambia su conformación en presencia de concentraciones subinhibitorias de ciertos antibióticos, modificando su afinidad por los operadores de DNA a los que se une (Ramos *et al.*, 2005; Jeng *et al.*, 2008; Chang *et al.*, 2010). En el caso de IcaR, se ha demostrado que la formación de dímeros sólo se inhibe por la acción de la Gentamicina (Jeng *et al.*, 2008).

Además, existen otros reguladores transcripcionales globales como SarA o SigB que han sido implicados en la regulación del biofilm dependiente de PIA-PNAG (Rachid *et al.*, 2000; Valle *et al.*, 2003; Tormo *et al.*, 2005). Sin embargo los mecanismos moleculares implicados en estos procesos de regulación no han sido aun explicados. Teniendo en cuenta que en otras bacterias el control del proceso de formación de biofilm ocurre a todos los niveles (transcripcional, post-transcripcional mediado por RNA reguladores y post-traducciona mediado tanto por procesamiento proteolítico como por factores alostéricos), es esperable que en *S. aureus* la regulación de la síntesis de PIA-PNAG también incluya elementos adicionales a los reguladores de transcripción descritos hasta ahora.

OBJETIVOS

2. OBJETIVOS DE LA TESIS DOCTORAL

Esta Tesis Doctoral forma parte de una línea de investigación dedicada al estudio de los mecanismos de regulación post-transcripcionales mediados por moléculas de RNAs y sus proteínas asociadas. Esta línea se inició con la determinación del mapa transcriptómico de alta resolución de *Staphylococcus aureus* que ha permitido la identificación de los inicios y finales de transcripción de todos los transcritos expresados en esta bacteria. Los análisis preliminares revelaron que una gran parte del transcriptoma de esta bacteria corresponde a regiones no codificantes, incluidas las regiones no traducidas del RNA mensajero. Teniendo en cuenta esta información, los objetivos que se plantearon para el desarrollo de esta Tesis doctoral fueron los siguientes:

- A) Realizar un análisis global y exhaustivo del transcriptoma de *Staphylococcus aureus* con el fin de identificar las regiones no traducidas de los RNA mensajeros.
- B) Realizar un análisis comparativo de las regiones 3' no traducidas del RNA mensajero para identificar aquellas con potencial para albergar elementos reguladores.
- C) Caracterizar a nivel molecular la función de las regiones 3' no traducidas (seleccionadas en el objetivo anterior) que pertenezcan a genes importantes para la fisiología de *S. aureus*.

MATERIALES Y MÉTODOS

3. MATERIALES Y MÉTODOS

3.1. Oligonucleotides, plasmids, bacterial strains and culture conditions

Bacterial strains, plasmids and oligonucleotides used in this study are listed in Table 2, Table 3 and Table 4 respectively (please see at the end of this section). *Staphylococcus aureus* strains were grown in trypticase soy broth supplemented with 0.25% glucose (TSB-gluc) (Pronadisa). *Escherichia coli* was grown in LB broth (Pronadisa). When required for selective growth, medium was supplemented with appropriated antibiotics at the following concentrations: erythromycin (Em), 1.5 $\mu\text{g ml}^{-1}$ and 20 $\mu\text{g ml}^{-1}$; ampicillin (Amp), 100 $\mu\text{g ml}^{-1}$.

3.2. RNA extractions

Bacteria were grown in 20 ml of TSB-gluc at 37°C under shaking conditions (200 rpm) until the culture reached an OD_{600nm} of 0.8. Cultures were centrifuged, the pellets were frozen in liquid nitrogen and stored at -80°C until needed. Total RNA from bacterial pellets was extracted using the TRIzol reagent method as described (Toledo-Arana *et al.*, 2009). Briefly, bacterial pellets were resuspended into 400 μl of solution A (glucose 10%, Tris 12.5 mM, pH 7.6, EDTA 10 mM) and mixed with 60 μl of 0.5 M EDTA. Resuspended cells were transferred into Lysing Matrix B tubes (MP Biomedicals) containing 500 μl of acid phenol pH 4.5 (Ambion) and mixed.

Bacteria were mechanically lysed with a Fastprep apparatus (BIO101) at speed 6.0 during 45 s at 4 °C. After lysis, tubes were centrifuged for 10 min at 17,900 g at 4 °C. The aqueous phase was transferred to 2-ml tubes containing 1 ml of TRIzol (Invitrogen), mixed, and incubated for 5 min at room temperature. 100 µl of chloroform were added, mixed gently, and incubated for 3 min at room temperature. Tubes were centrifuged for 10 min at 17,900 g at 4°C. The aqueous phase was transferred into a 2-mL tube containing 200 µl of chloroform, mixed, and incubated for 5 min at room temperature. Tubes were centrifuged for 5 min at 17,900 g at 4 °C. RNA contained in the aqueous phase was precipitated by addition of 500 µl of isopropanol and incubated for 15 min at room temperature. Tubes were centrifuged for 15 min at 17,900 g at 4 °C. RNA pellets were washed with 75% ethanol. Dried RNA pellets were resuspended in DEPC-treated water. RNA concentrations were quantified, and RNA qualities were determined by using Agilent RNA Nano LabChips (Agilent Technologies). RNAs were stored at –80 °C until needed.

3.3. RNA sequencing

Total RNA was extracted after lysing bacteria in the Fastprep homogenizer (Bio101) using TRIzol reagent (Invitrogen). The RNA fraction shorter than 50 nt was purified from total RNA (100 mg) through electrophoresis at 75V constant voltage during 16 min with flashPAGE fractionator (AMBION). Purified RNA fraction was incubated with 2U of TAP (Epicenter) for 2h at 37°C to convert 5' triphosphate groups to monophosphates. The TAP

treated short RNA fraction was ligated to the 3' ModBan oligonucleotide (Table 4) by using the truncated T4 RNA ligase². This was followed by a second gel purification on 10% TBE-UREA gel to remove the non-ligated 3' ModBan oligonucleotide. The extracted product was ligated to 5' RNA Illumina adapter by using T4 RNA ligase. cDNA was produced using SuperScript III and the RT oligonucleotide, complementary to the ModBan 3' linker. The resulting RT product was then amplified using Sol_5_SBS3 and Sol_3_ModBan adapters, purified and sent for Illumina sequencing.

3.4. cDNA synthesis, labelling and tiling array hybridization

Before cDNA synthesis, RNA integrity from each sample was confirmed on Agilent RNA Nano LabChips (Agilent Technologies). 10 µg of total RNA were reverse transcribed using SuperScript II reverse transcriptase (Invitrogen Life Technologies) and processed following the protocol of the Affymetrix GeneChip Expression Analysis Technical Manual (P/N 702232 Rev. 2) in the presence of 6 ng/ml Actinomycin D to avoid spurious second-strand cDNA synthesis during reverse transcription reaction (Perocchi *et al.*, 2007). Sense RNA corresponding to *B. subtilis* poly-A *lys, phe, thr, trp, dap* genes were spiked into sample RNA as control for labelling and hybridization steps. cDNA was digested by DNase I (PIERCE) in 10X DNase I buffer (USB-Affymetrix) and the size of digestion products was analysed in the Agilent Bioanalyser 2100 using RNA Nano LabChips to ensure that the fragmentation resulted in a majority of products in the range of 50 to 200 base-pairs. The fragmented cDNA was then biotinylated using

terminal deoxynucleotidyl transferase (Promega) and the GeneChip DNA labelling reagent (Affymetrix) following the manufacturer's recommendations. Biotinylated cDNA (5 microgram per array) was hybridized on custom *S. aureus* tiling microarrays designed as described (Segura *et al.*, 2012). Hybridization was carried out during 16 h according to the Affymetrix protocol in a total volume of 200 µl per hybridization chamber. Following incubation, the arrays were washed and stained in the Fluidics station 450 (Affymetrix) using the protocol n°FS450_0005. Scanning of the arrays was then performed using the GeneChip scanner 3000 (Affymetrix). A first scan of the chip was carried out with gene expression sub-array parameters followed by a second scan with tiling sub-array parameters. Intensity signals of each probe cells were computed by the GeneChip operating software (GCOS) and stored in cell intensity files (.CEL extension) before preprocessing and analysis.

3.5. Microarray data analysis

Data analysis of the tiling sub-array was performed using the Tiling Analysis Software (TAS) from Affymetrix (<http://www.affymetrix.com>). Output bar files containing probe signal values were converted in graphic type files (.gr extension file) to be loaded at the *Staphylococcus aureus* transcriptome browser (<http://staph.unavarra.es/>).

3.6. cDNA libraries for RNA-seq, read mapping and statistics analysis

Deep sequencing of RNAs from *S. aureus* 15981 strain was performed as previously described (Lasa *et al.*, 2011). Mapped reads were included in .wig files to be loaded at the *Staphylococcus aureus* transcriptome browser (<http://staph.unavarra.es/>).

3.7. Simultaneous mapping of 5'- and 3'-ends of RNA molecules

The simultaneous mapping of 5'- and 3'-ends of the entire *icaR* mRNA molecule and the processing sites was performed by RACE (Rapid Amplification of cDNA Ends) using circularized RNAs as previously described (Britton *et al.*, 2007), with the following modifications. Specifically, reactions were performed on RNAs extracted from bacteria grown in TSB-gluc until an OD_{600nm} = 0.8 was reached. Six µg of RNA were treated with TURBO DNase I (Ambion). After phenol extraction to inactivate DNase I, the RNA was divided into two aliquots. Both aliquots were incubated for 45 min at 37°C with the corresponding buffer, in the presence or absence of Tobacco Acid Pyrophosphatase, (TAP) (Epicentre biotechnologies) respectively. This step allows discriminating a 5'-end generated by transcription initiation from a 5'-end provided by RNA processing. The TAP treatment step was avoided when processing sites wanted to be determined. After incubation, acid-phenol and chloroform extractions and ethanol precipitation was performed. Serial dilutions (from 500 ng to 0.5 ng) of the TAP+ and TAP- treated RNAs were prepared. Each dilution was ligated with 40 U of T4 RNA ligase I (New England Biolabs) in the presence

of 1X RNA ligase Buffer, 8% DMSO, 10 U of RNase Inhibitor, 1 U of DNase I and RNase-free water in a total volume of 25 µl at 17°C overnight. After acid-phenol and chloroform extractions and ethanol precipitation, the ligated RNAs were resuspended in 10 µl of RNase-free water. RT-PCR reactions were performed using specific outward primers (Table 4) and the SuperScript One-Step RTPCR kit (Invitrogen). RT-PCR products were run on 3% TAE-agarose gels. For mapping of the entire size of the molecule, bands only present in the TAP+ reactions were purified by Gel Extraction kit (QIAGEN) and cloned using TOPO TA Cloning kit (Invitrogen). For mapping of processing sites, all bands observed in the gel were purified and cloned. Eight transformants per cloned band were analysed by PCR using M13 forward and reverse primers. Plasmids containing the expected insert size were sent to sequencing. To determine the localization of the 5'- and 3'-ends, sequences were compared with *icaR* region from the *S. aureus* 132 genome sequence (Vergara-Irigaray *et al.*, 2009).

3.8. Riboprobes synthesis

Strand-specific riboprobes to detect *icaR* mRNA were synthesized from a PCR product containing a T7 phage promoter sequence (see Table 4 for oligonucleotides). One microgram of these PCR products was used as a matrix template for *in vitro* transcription reaction with phage T7 RNA polymerase, 0.5 mM each ATP, GTP, CTP, and 50 mCi of [α -32P] UTP using the Maxiscript kit (Ambion). The riboprobes were then treated with TURBO DNase I at 37°C for 30 min, and reactions were stopped by addition

of 1 µl of 0.5M EDTA. The riboprobes were purified on Bio-Spin 30 columns following the manufacturer's recommendations (Bio-Rad) and were immediately used.

3.9. Northern blots

Northern blots were performed as described (Toledo-Arana *et al.*, 2009). Briefly, 8-15 µg of total RNA were separated in precast 1.25% agarose gels (Sigma) by using 1X NorthernMax MOPS as running buffer (Ambion). After electrophoresis, gels were stained with ethidium bromide and photographed to verify equal loading of RNA samples. Then, RNAs were transferred onto Nytran membranes (0.2 µm pore size) (Sigma) by using NorthernMax One Hour Transfer buffer reagent as described in the manufacturer's protocol (Ambion). RNA was UV cross-linked to the membrane by using the UV Stratalinker 1800 (Stratagene). Membranes were prehybridized for at least 30 min in ULTRAhyb solution (Ambion) at 65°C, followed by addition of labelled strand-specific riboprobe and overnight hybridization at 65 °C. Membranes were then washed twice with 2X SSC-0.1% SDS for 5 min at 65 °C. The size of the transcripts was estimated by comparison with RNA Millennium molecular weight standards (Ambion). Autoradiography images were registered at different exposition times according to each experiment.

3.10. Chromosomal allelic exchange

To generate the *icaR* 3'-UTR and *dpbA* deletions, we amplified by PCR two fragments of approximately 500 pb that flanked the left (primers A and B, Table 4) and right sequences (primers C and D, Table 4) of the region targeted for deletion. The PCR products were amplified with Phusion® High-Fidelity DNA Polymerase (Fermentas-Thermo Scientific), purified and cloned separately in pCR®-Blunt II TOPO vector (Invitrogen). Fragments were then fused by ligation into the shuttle vector pMAD (Arnaud *et al.*, 2004). The resulting plasmid was transformed into *S. aureus* 15981 or 132 strains by electroporation. Homologous recombination experiments were performed as described (Valle *et al.*, 2003). Erythromycin sensitive white colonies, which no longer contained the pMAD plasmid, were tested by PCR using primers E and F and DNA sequencing.

3.11. Plasmid constructions

All PCR fragments were amplified from *S. aureus* 132 chromosomal DNA using Phusion® High-Fidelity DNA Polymerase (Fermentas-Thermo Scientific) using the appropriate oligonucleotides. PCR fragments were purified and cloned into pCR®-Blunt II TOPO vector (Invitrogen). DNA fragments were excised from this vector with appropriate restriction enzymes and then subcloned into the shuttle vectors pSA14 (Joanne *et al.*, 2009) or pCN40 (Charpentier *et al.*, 2004). pSA14 plasmid is a pMK4 derivative carrying promoterless *E. coli lacZ* gene for constructing

transcriptional fusions while pCN40 plasmid allows the expression of the gene of interest from the constitutive P_{blaZ} promoter. To construct pSA14-Pica plasmid, a 422 bp DNA fragment, which includes the *ica* operon promoter, was amplified with *icaA*-pSA14-Fw and *icaA*-pSA14-Rv oligonucleotides. pCN40 plasmid derivatives, which expressed different *icaR* mRNA versions from the constitutive P_{blaZ} promoter, were constructed as follows. *plcaRm* plasmid was constructed by amplifying a PCR fragment of 1,079 nt with *IcaR*+1 and *IcaR*-Term oligonucleotides and cloning into the BamHI/EcoRI site of pCN40. *plcaRm* Δ 3'UTR plasmid was constructed by amplifying a PCR fragment of 748 nt with *IcaR*+1 and *IcaR*-Term oligonucleotides using *S. aureus* 132 Δ 3'-UTR chromosomal DNA as template and cloning into BamHI/EcoRI sites of pCN40. To construct p^{FLAG} *IcaRm*_WT, p^{FLAG} *IcaRm* Δ 3'UTR, *plcaRm* Δ anti-SD, *plcaRm*_SUBST, p^{FLAG} *IcaRm* Δ anti-SD, p^{FLAG} *IcaRm*_SUBST and *plcaRm*-Compensatory plasmids, overlapping PCR was used with oligonucleotides shown in Table 4. All constructed plasmids were confirmed by sequencing.

3.12. Quantitative reverse transcription PCR

Total RNA from bacterial cells grown until OD_{600nm} of 0.8 was extracted as described above. Each RNA sample was subjected to TURBO DNase I (Ambion) treatment for 30 min at 37°C. The enzyme was inactivated by phenol-chloroform extractions. RNA quality was assessed with an Agilent 2100 Bioanalyser. Twenty μ l of random primers (50 ng μ l⁻¹; Invitrogen) and 20 μ l of deoxynucleoside triphosphates (dNTPs) (10 mM mix; Invitrogen)

were added to the samples containing 8 to 10 μg of RNA in a volume of 100 μl of diethyl pyrocarbonate (DEPC) water. After 5 min of incubation at 65°C, samples were chilled on ice at least during 1 min, and a reverse transcription (RT) mix containing 44 μl of 5X first-strand buffer (Invitrogen), 22 μl of dithiothreitol (DTT) (0.1 M; Invitrogen), 2 μl of SuperScript III Retrotranscriptase (200 U μl^{-1} ; Invitrogen) and 1 μl of RNase Out (40 U μl^{-1} ; Invitrogen) was added to each preparation. cDNA was obtained after a cycle of 10 min at 25°C, 50 min at 50°C, and 5 min at 85°C. RNA was eliminated by the addition of 1 μl of RNase H (10 U μl^{-1} ; Invitrogen) and incubation for 20 min at 37°C. cDNA samples were purified with CentriSep spin columns (Princeton separations). cDNA concentration was adjusted to 100 ng μl^{-1} . One μl of the cDNA samples was used for real-time quantitative PCR using SYBR green PCR master mix (Applied Biosystems) and the ABI Prism 7900 HT instrument (Applied Biosystems). The PCR was performed under the following conditions: 95°C for 20 s, 40 cycles of 95°C for 1 s and 60°C for 20 s, and a final step at 95°C for 15 s, 60°C for 15 s, and 95°C for 15 s. *icaR* and *gyrB* mRNA levels were quantified by cDNA amplification using oligonucleotides described in Table 4 and values were normalized to those of the housekeeping *gyrB* gene.

3.13. Western blot analysis

Overnight cultures of the strains tested were diluted 1:100 in TSB-gluc, and 20 ml of this cell suspension were grown at 37 or 23 °C in 125 ml flasks until

OD_{600nm} reached 0.8. Ten ml of bacterial cultures were centrifuged and pellets were resuspended in 100 μ l PBS. Then, 2 μ l of Lysostaphin 1 mg ml⁻¹ (Sigma) and 3 μ l of DNase I 1mg ml⁻¹ (Sigma) were added. After 2 h of incubation at 37°C cell lysates were centrifuged and supernatants were collected. Protein concentration was determined with the Bio-Rad protein assay (Bio-Rad). Samples were adjusted to 5-10 μ g μ l⁻¹ of total protein and one volume of Laemmli buffer was added. Total protein extracts were denatured by boiling at 100°C for 5 min. Proteins were separated on 12% SDS-polyacrylamide gels and stained with 0.25% Coomassie brilliant blue R250 (Sigma) as loading controls. For Western blotting, proteins were transferred onto Hybond-ECL nitrocellulose membranes (Amersham Biosciences) by semi-dry electroblotting. Membranes were blocked overnight with 5% skimmed milk in phosphate-buffered saline (PBS) with 0.1% Tween 20, and incubated with anti-FLAG antibodies labelled with phosphatase alkaline (Sigma) diluted 1:500 for 2 h at room temperature. 3XFLAG labelled IcaR protein was detected with the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

3.14. mRNA stability assays

Overnight cultures were diluted 1:100 in TSB-gluc, and 150 ml of these cell suspensions were grown in 500 ml flasks until OD_{600nm}=0.8 was reached. Twelve ml of cultures were transferred to six sterile 15 ml falcon tubes containing 300 μ g ml⁻¹ Rifampicin. The tube corresponding to time 0 min also contained 2.5 mL STOP solution (5% phenol equilibrated at pH 7 and

95% of ethanol). After addition of the culture, the time “0 min” tube was immediately centrifuged at 4,500 rpm during 3 min, the supernatant discarded and the pellet frozen in liquid nitrogen. The rest of the tubes were incubated at 37°C and 2.5 ml of STOP solution were added at times 2, 4, 8, 15 and 30 min respectively after rifampicin addition. Then, each tube was centrifuged at 4,500 rpm during 3 min. Supernatants were discarded and pellets were frozen in liquid nitrogen and stored at -80°C until needed. RNA extractions and Northern blots were performed as described above.

3.15. mRNA dimerization assays on native agarose gel electrophoresis

The full-length *icaR* mRNA was produced by *in vitro* transcription using T7 RNA polymerase. To visualize the monomeric form of the mRNA, *icaR* mRNA (0.25 µg) was denatured 3 min at 90°C in 10 µl of sterile bi-distillated water or of a buffer containing Tris-HCl 20 mM pH 7.5, 1 mM EDTA, chilled on ice for 1 min followed by an incubation at 37°C for 15 min. To visualize alternative conformations of *icaR* mRNA, the mRNA (0.25 µg) was first denatured in 8 µl of sterile bi-distillated water for 3 min at 90°C, chilled on ice, and renatured at 37°C for 15 min by adding 2 µl of a 5 x concentrated buffer containing Tris-HCl 100 mM pH 7.5, 250 mM KCl in the absence or in the presence of 50 mM MgCl₂. All samples were then mixed with 2 µl of loading buffer (48% glycerol, 0.01% bromophenol blue) and electrophoresed on 1% agarose gel in 0.5X TBE buffer. The RNA was then visualized after ethidium bromide staining.

3.16. Gel shift assays

5'-UTR and 3'-UTR wild type PCR fragments (117 and 120 nt respectively) used as templates for T7 *in vitro* transcription were amplified using chromosomal DNA from *S. aureus* 132 strain, while 5'-UTR-compensatory and 3'-UTR-substituted fragments were amplified from the corresponding plasmids using oligonucleotides shown in Table 4. RNA molecules used as substrates in gel shift assays were synthesized by *in vitro* T7 transcription using Riboprobe *in vitro* Transcription System (Promega). When needed 50 mCi of [α -³²P] UTP was used for radiolabelling. RNA fragments were purified by electrophoresis on an 8.3 M urea/6% polyacrylamide gel. The bands were excised from the gel and RNA was eluted with elution buffer (3 M ammonium acetate pH 5.2, 1mM EDTA, 2.5% (v/v) phenol pH 4.3) overnight at room temperature. Then, RNA was ethanol precipitated, resuspended in RNase free water and quantified using the Biophotometer Plus (Eppendorf). The yield of the labelled substrates (cpm μ l⁻¹) was determined by scintillation counting.

Binding assays were performed in 1X TMN buffer (20 mM Tris acetate pH7.6, 100 mM sodium acetate, 5 mM magnesium acetate) as previously described (Udekwu *et al.*, 2005). Briefly, labelled RNA molecules (0.025 pmol of 5'-UTR-WT or 5'-UTR-compensatory) were incubated with increasing concentrations of unlabelled RNA fragments (3'-UTR WT or 3'-UTR-substituted) in a total volume of 10 μ l, at 37°C or 23°C for 30 min. Binding reactions were then mixed with 2 μ l of loading buffer (48% glycerol, 0.01% bromophenol blue) and electrophoresed on native 5%

polyacrylamide gels in 0.5X TBE buffer at 200 V in a cold room for 3 h. Gels were analysed using a PhosphorImager (Molecular Dynamics).

3.17. Purification of recombinant *S. aureus* RNase III

S. aureus RNase III coding sequence (*rnc*) was amplified with primer pairs RNase III Fw (NdeI) and RNase III Rv (BamHI). The purified PCR fragment was double digested with BamHI and NdeI and ligated into the pET-15b vector (Promega), generating plasmid pET-15b RNase III. This plasmid was transformed in *E. coli* BL21(DE3) *rnc105 recA*. This strain is slow growing but allows overproduction of His₆-RNase III (Amarasinghe *et al.*, 2001). The BL21(DE3) *rnc105 recA* strain carrying the pET-15b RNase III plasmid was grown in 200 ml of LB medium supplemented with ampicillin (100 µg ml⁻¹) to an OD_{600nm} of 1.5. At this point, protein expression was induced by addition of 1 mM IPTG and the culture was further incubated overnight at 15°C. Cells were then harvested by centrifugation, the pellet was washed with 12 ml of cold buffer (25 mM Tris-HCl pH 8.0, 8% ammonium sulphate, 0.1 mM EDTA) and resuspended in 6 ml of the same buffer. Cells were lysed using a French Press at 900 psi in the presence of 0.1 mM of PMSF. After lysis, the crude extracts were treated with 125 U of Benzonase (Sigma). The protein extract was then clarified by centrifugation for 30 min, 27,000g at 4°C. The histidine tagged recombinant RNase III was purified by affinity chromatography, using the ÄKTA FPLCTM System (GE Healthcare). The clarified extracts were loaded into a HisTrap HP Sepharose 1 ml column. Protein elution was achieved with a linear imidazole gradient (from 0 mM to

300 mM) in buffer B (25 mM Tris-HCl pH 8.0, 1M NH₄Cl, 300 mM imidazole). Fractions containing the protein of interest, free of contaminants, were pooled and buffer exchanged by dialysis against Desalting Buffer (25 mM Tris-HCl pH 8.0, 500 mM KCl, 0,1 mM DTT, 50 % glycerol) using a Slide-A-Lyzer Dialysis Cassette (Thermo Scientific) with a molecular mass cut-off of 10 kDa. Protein samples were quantified using the Bio-Rad protein assay (Bio-Rad) and stored at -20°C. The purity of the enzyme was analysed by SDS-PAGE.

3.18. *In vitro* RNase III activity assays

RNA substrates were prepared as described in the Gel shift assays section. The activity of purified RNase III was tested over [³²P]- α -UTP labelled 5'-UTR-WT fragment in the presence or absence of 3'-UTR-WT or 3'-UTR-substituted fragments. Hybridization between labelled and unlabelled substrates was performed in a 1:50 molar ratio in the Tris component of the activity buffer by incubation for 10 minutes at 80°C, followed by 45 minutes at 37°C or 23°C. The same treatment was applied to free 5'-UTR-WT. Activity assays were carried out in a final volume of 40 μ l containing the activity buffer (30 mM Tris-HCl pH 8, 160 mM NaCl and 0.1 mM DTT) and approximately 0.14 pmol of substrate. 10 mM MgCl₂ were added to the reaction mixture. As a control, prior to the beginning of each assay, an aliquot (without the enzyme) was taken and incubated in the same conditions until the end of the assay. Reactions were started by the addition of the enzyme at a concentration of 500 nM and were incubated at 37°C or

23°C (Chelladurai *et al.*, 1991). Samples were withdrawn at different times and reactions were stopped by the addition of formamide-containing dye supplemented with 10 mM EDTA. Reaction products were run in a 7M urea/10% polyacrylamide gel, visualized by PhosphorImaging and analysed using ImageQuant software (Molecular Dynamics).

3.19. Toeprinting assays

Full length *icaR* mRNA or its variants were cloned (Serganov *et al.*, 1997) into *Stu*I and *Bam*HI restriction sites of pUT7 plasmid for *in vitro* transcription. The whole 5'-UTR including 72-nt of the ORF and the 3'-UTR (186 nt) were transcribed from PCR products obtained with oligos T7 5'-UTR and *IcaR* 5' rev, and T7 3'-UTR and *IcaR* 3' rev, respectively (Table 4). *S. aureus* 30S ribosomal subunits were prepared as described (Fechter *et al.*, 2009). The formation of a simplified translation initiation complex with mRNAs and the extension conditions were as described (Fechter *et al.*, 2009). Standard conditions contained 5 nM *icaR* mRNA or *icaR* 5'-UTR annealed to a 5' end labelled oligonucleotide (*IcaR* 5' rev), 500 nM *S. aureus* 30S ribosomal subunits, and 0.5 to 10 µM of *icaR* 3'-UTR in 10 µl of buffer containing 20 mM Tris-acetate pH 7.5, 60 mM NH₄Cl, 8.5 mM magnesium acetate and 1 mM DTT. The 30S subunits were renatured for 10 min at 37°C before incubation with the mRNAs. After 10 min of incubation at 37°C, the initiator tRNA (1µM) was added, and the reaction was further incubated for 5 min at 37°C. Reverse transcription was conducted with one unit of AMV reverse transcriptase for 15 min at 15°C.

Toeprint was run on 8% polyacrylamide gels and visualized by autoradiography.

3.20. PIA-PNAG quantification

Cell surface PIA/PNAG exopolysaccharide levels were quantified as previously described (Cramton *et al.*, 1999). Briefly, overnight cultures of the strains tested were diluted 1:40 in the appropriate medium and 2 ml of this cell suspension were used to inoculate sterile 24-well polystyrene microtiter plates (Sarstedt). After 24 h of static incubation at 37°C or 23°C, the same number of cells of each strain was resuspended in 50 μ l of 0.5 M EDTA (pH 8.0). Then, cells were incubated for 5 min at 100°C and centrifuged 17,000 g for 5 min. Each supernatant (40 μ l) was incubated with 10 μ l of proteinase K (20 mg ml⁻¹) (Sigma) for 30 min at 37°C. After the addition of 10 μ l of Tris-buffered saline (20 mM Tris-HCl, 150 mM NaCl [pH 7.4]) containing 0.01% bromophenol blue, 5 μ l were spotted on a nitrocellulose membrane using a Bio-Dot microfiltration apparatus (Bio-Rad). The membrane was blocked overnight with 5% skimmed milk in phosphate-buffered saline (PBS) with 0.1% Tween 20, and incubated for 2 h with specific anti-PNAG antibodies diluted 1:10,000 (Maira-Litrán *et al.*, 2005). Bound antibodies were detected with peroxidase-conjugated goat anti-rabbit immunoglobulin G antibodies (Jackson ImmunoResearch Laboratories, Inc., Westgrove, PA) diluted 1:10,000 and developed using the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

3.21. Biofilm formation assay

Biofilm formation assay in microtiter plates was performed as described (Heilmann *et al.*, 1996). Briefly, strains were grown overnight at 37°C or 23°C and were diluted 1:40 in TSB-gluc and TSB-NaCl for *S. aureus* 15981. 200 µl of cell suspension was used to inoculate sterile 96-well polystyrene microtiter plates (Iwaki). After 24 hours at 37°C or 23°C, wells were gently rinsed by immersion in a water bath three times. Bacteria adhered at the bottom of the well was developed by staining with 0.1% of crystal violet for 5 min. After incubation, wells were gently rinsed three times by immersion in a water bath and dried before photographed. Then, crystal violet adhered at the bottom of the wells was resuspended with 200 µl of a solution of ethanol/acetone (80/20) and quantified using a Multiskan EX with a 595 nm filter.

To analyse biofilm formation under flow conditions, we used 60-ml microfermentors (Ghigo, 2001) (Pasteur Institute; (www.pasteur.fr/recherche/unites/Ggb/biofilmfermenter.html) with a continuous flow of 40 ml of TSB h⁻¹ and constant aeration with sterile pressed air as previously described (Valle *et al.*, 2003). Submerged Pyrex slides served as the growth substratum. 10⁸ bacteria from an overnight preculture grown in TSB-gluc of each strain were used to inoculate microfermenters and were cultivated 8 h at 37°C. Biofilm development was recorded with a digital camera.

3.22. RNA structural probing

Enzymatic hydrolysis and chemical modifications on free or bound RNAs were performed as described previously (Benito *et al.*, 2000). Briefly, 5'-UTR, 3'-UTR and *icaR* mRNA full length were transcribed *in vitro* with T7 RNA polymerase as explained in Toeprint assay. Enzymatic hydrolysis was performed on 5'-end radioactively labelled substrates after 30 minutes of refolding at 23°C or 37°C. The 5'-UTR alone or in conjunction with the 3'-UTR was probed with RNase T1 (unpaired G), RNase V1 (double stranded RNA), and RNase T2 (unpaired A) and DMS (N3A, N1A) in a buffer A (Hepes-KOH 50 mM pH 7.5/MgCl₂ 1 or 5 mM/ KCl 50 mM). Alkylation of C(N3) and A(N1) positions was done with 1 µl DMS (Tijerina *et al.*, 2007) (diluted 1/8 and 1/16 in ethanol) for 2 min, and modification of U(N3) and G(N1) were done with 5 µl of CMCT (50 mg/ml) for 10, 15 and 20 min in a buffer A. Unpaired riboses were tested with 5 mM of Shape reagent for 5, 10, 15 and 20 minutes (Edward J Merino *et al.*, 2005; Lusvarghi *et al.*, 2012). After chemical modifications, the substrates were subjected to primer extension with reverse transcriptase with a 5' radioactively labelled oligonucleotide. The reactions were stopped with 5 µl of EDTA 0.1 M for induced cleavages and all other reactions were stopped by phenol extraction followed by RNA precipitation. The 5'-end-labelled fragments were electrophoresed on 6%-12% polyacrylamide/8 M urea slab gels. The cleavage positions were identified by running sequencing reactions and alkaline ladders.

Table 2. Strains used in this study.

| Strains | Relevant characteristic(s) | MIC ^a | Source or reference ^b |
|--|--|------------------|---|
| <i>Staphylococcus aureus</i> | | | |
| 15981 | Clinical strain. Biofilm positive in TSBg | 532 | (Valle <i>et al.</i> , 2003) |
| MW2 | Typical community-acquired strain of MRSA, which was isolated in 1998 in North Dakota, USA. | 3566 | NRS123, (Baba <i>et al.</i> , 2002) |
| RN10359 | RN450 lysogenic for 80 α phage | 3337 | (Úbeda <i>et al.</i> , 2007) |
| ISP479r | ISP479c with <i>rsbU</i> gene restored | 1680 | (Toledo-Arana <i>et al.</i> , 2005) |
| 132 | MRSA clinical strain with the capacity to form proteinaceous biofilm in TSB-glucose and PIA-PNAG-dependent biofilm in TSB-NaCl | 29 | (Vergara-Irigaray <i>et al.</i> , 2009) |
| 15981 Δ 3'-UTR | 15981 strain with deletion in the <i>icaR</i> 3' UTR | 2695 | This study |
| 132 Δ 3'-UTR | 132 strain with deletion in the <i>icaR</i> 3' UTR | 3161 | This study |
| 15981 ^{FLAG} IcaR | 15981 strain expressing the 3XFLAG tagged IcaR protein from the chromosome | 2653 | This study |
| 15981 ^{FLAG} IcaR Δ 3'-UTR | 15981 Δ 3'-UTR strain expressing the 3XFLAG tagged IcaR protein from the chromosome | 2649 | This study |
| RN9594 | RN4220 strain carrying the pCN40 plasmid | 2534 | NRS594 |
| 15981 pCN40 | 15981 strain harbouring the pCN40 plasmid | 1014 | This study |
| 15981 p ^{FLAG} IcaRm | 15981 strain harbouring the p ^{FLAG} IcaRm plasmid | 2683 | This study |
| 15981 p ^{FLAG} IcaRm Δ 3'UTR | 15981 strain harbouring the p ^{FLAG} IcaRm Δ 3'UTR plasmid | 2684 | This study |
| 15981 p ^{FLAG} IcaRm Δ anti-SD | 15981 strain harbouring the p ^{FLAG} IcaRm Δ anti-SD plasmid | 4392 | This study |
| 15981 p ^{FLAG} IcaRm_SUBST | 15981 strain harbouring the p ^{FLAG} IcaRm-SUBST-anti-SD plasmid | 4395 | This study |
| 15981 pIcaRm | 15981 strain harbouring the pIcaRm plasmid | 1011 | This study |
| 15981 pIcaRm Δ 3'UTR | 15981 strain harbouring the pIcaRm Δ 3'UTR plasmid | 1012 | This study |
| 15981 pIcaRm Δ anti-SD | 15981 strain harbouring the pIcaRm Δ anti-SD plasmid | 4391 | This study |
| 15981 pIcaRm_SUBST | 15981 strain harbouring the pIcaRm-SUBST-anti-SD plasmid | 4394 | This study |
| 15981 Δ <i>rnc::cat86</i> | 15981 strain with deletion in the <i>rnc</i> gene | 1327 | (Lasa <i>et al.</i> , 2011) |
| 15981 Δ <i>pnp</i> | 15981 strain with deletion in the <i>pnp</i> gene | 1100 | (Lasa <i>et al.</i> , 2011) |
| 15981 Δ <i>yqfR</i> | 15981 strain with deletion in the SA1387 gene which encodes a DEAD-box RNA helicase | 1394 | This study |
| 15981 Δ <i>hfq::cat86</i> | 15981 strain with deletion in the <i>hfq</i> gene | 1055 | This study |
| 15981 Δ <i>rnc</i> p ^{FLAG} IcaRm | 15981 Δ <i>rnc::cat86</i> strain harbouring the p ^{FLAG} IcaRm plasmid | 2706 | This study |
| 15981 Δ <i>rnc</i> p ^{FLAG} IcaRm Δ 3'UTR | 15981 Δ <i>rnc::cat86</i> strain harbouring the p ^{FLAG} IcaRm Δ 3'UTR plasmid | 2707 | This study |
| 15981 Δ <i>rnc</i> p ^{FLAG} IcaRm_SUBST | 15981 Δ <i>rnc::cat86</i> strain harbouring the p ^{FLAG} IcaRm-SUBST-anti-SD plasmid | 4970 | This study |
| 15981 Δ <i>yqfR</i> p ^{FLAG} IcaRm | 15981 Δ <i>yqfR</i> strain harbouring the p ^{FLAG} IcaRm plasmid | 2704 | This study |
| 15981 Δ <i>hfq</i> p ^{FLAG} IcaRm | 15981 Δ <i>hfq::cat86</i> strain harbouring the p ^{FLAG} IcaRm plasmid | 2708 | This study |
| 15981 pSA14-PicaA53 | 15981 strain harbouring the pSA14-PicaA53 plasmid | 1449 | This study |
| <i>Escherichia coli</i> strains | | | |
| XL1 Blue | Used for cloning experiments | | Stratagene |
| BL21(DE3) <i>rnc105</i> | BL21(DE3) <i>mc105 recA</i> | 4848 | (Amarasinghe <i>et al.</i> , 2001) |
| BL21 pET-15b RNase III | BL21(DE3) <i>rnc105 recA</i> strain carrying the pET-15b expressing the <i>S. aureus</i> RNase III fused to His-tag | 4867 | This study |

^aMicrobial Biofilm Laboratory strain collection number.

Table 3. Plasmids used in this study.

| Plasmids | Relevant characteristic(s) | Source or reference ^a |
|---------------------------------|--|------------------------------------|
| pMAD | <i>E. coli</i> - <i>S. aureus</i> shuttle vector with a thermosensitive origin of replication for Gram-positive bacteria. The vector contains the <i>bgaB</i> gene encoding a b- galactosidase under the control of a constitutive promoter as reporter of plasmid presence. Ap ^R , Em ^R . | (Arnaud <i>et al.</i> , 2004) |
| pMAD Δ3'UTR | pMAD plasmid containing the mutant allele for deletion of the long 3'-UTR of the <i>icaR</i> mRNA | This study |
| pMAD 3xFLAG-icaR | pMAD plasmid containing the allele for insertion of the 3xFLAG at the N-terminus of the IcaR protein | This study |
| pMAD ΔSA1387 | pMAD plasmid containing the allele for deletion of the DEAD-box RNA helicase (<i>S. aureus</i> N315 ID: SA1387) | This study |
| pLUG533 | pMAD derivative for deletion/replacement of <i>S. aureus</i> <i>hfq</i> gene | (Boisset <i>et al.</i> , 2007) |
| pSA14 | <i>E. coli</i> - <i>S. aureus</i> shuttle vector for transcriptional LacZ reporter fusions. The plasmid carries the promoterless <i>Escherichia coli</i> <i>lacZ</i> gene downstream from the <i>Bacillus subtilis</i> <i>spoVG</i> ribosome binding site. | (Joanne <i>et al.</i> , 2009) |
| pSA14-Pica | pSA14 plasmid carrying the <i>ica</i> operon promoter region from -422 to +1, where +1 corresponds to the mapped transcriptional start site of <i>icaADBC</i> mRNA | This study |
| pCN40 | <i>E. coli</i> - <i>S. aureus</i> shuttle vector to express genes under the control of the P _{blaZ} constitutive promoter. Low copy number (20 to 25 copies/cell). Em ^R | (Charpentier <i>et al.</i> , 2004) |
| p ^{FLAG} IcaRm | pCN40 plasmid constitutively expressing the N-terminal 3XFLAG tagged IcaR protein from the entire mRNA | This study |
| p ^{FLAG} IcaRmΔ3'UTR | pCN40 plasmid constitutively expressing the N-terminal 3XFLAG tagged IcaR protein from the mRNA carrying a deletion of 331 nt downstream of the stop codon | This study |
| p ^{FLAG} IcaRmΔanti-SD | pCN40 plasmid constitutively expressing the N-terminal 3XFLAG tagged IcaR protein from the mRNA carrying the UCCCCUG deletion | This study |
| p ^{FLAG} IcaRm-SUBST | pCN40 plasmid constitutively expressing the N-terminal 3XFLAG tagged IcaR protein from the mRNA carrying the substitution of the UCCCCUG motif by AGGGGAC | This study |
| pIcaRm | pCN40 plasmid constitutively expressing the entire <i>icaR</i> mRNA molecule | This study |
| pIcaRmΔ3'UTR | pCN40 plasmid constitutively expressing the <i>icaR</i> mRNA carrying a deletion of 331 nt downstream of the stop codon | This study |
| pIcaRmΔanti-SD | pCN40 plasmid constitutively expressing the entire <i>icaR</i> mRNA carrying the UCCCCUG deletion | This study |
| pIcaRm-SUBST | pCN40 plasmid constitutively expressing the <i>icaR</i> mRNA carrying the substitution of the UCCCCUG motif by AGGGGAC | This study |
| pIcaRm-Compensatory | pCN40 plasmid constitutively expressing the <i>icaR</i> mRNA carrying the substitution of the UCCCCUG motif by AGGGGAC and the substitution of the SD region (CAGGGGG) by GTCCCCCT | This study |
| pET-15b RNase III | pET-15b expressing the <i>S. aureus</i> <i>rnc</i> gene which encodes the double stranded endoribonuclease RNase III fused to His-tag. | This study |
| pUT7- <i>spa</i> | T7 promoter- <i>spa</i> (nt -25 to +200) allowing the in vitro T7 transcription of <i>S. aureus</i> <i>spa</i> mRNA containing the whole 5'-UTR and 200 nt of the coding sequence | (Huntzinger <i>et al.</i> , 2005) |
| pUT7- <i>icaR</i> | T7 promoter- <i>icaR</i> allowing the in vitro T7 transcription of the full-length <i>icaR</i> mRNA | This study |

Table 4. Oligonucleotides used in this study.

| Oligonucleotide | Sequence |
|---|--|
| Small RNA library sequencing | |
| 3' ModBan | AMP-5'p-5'pCTGTAGGCACCATCAAT-ddC |
| 5' Illumina RNA adapter | OH-rArCrArCrUrCrUrUrCrCrUrArCrArCrGrArCrGrUrCrUrUrCrCrGrArUrC-OH |
| RT-PCR | ATTGATGGTGCCTACAG |
| Sol_5_SBS3 | AATGATACGGCGACCAACCAACTCTTCCCTACACGACG |
| Sol_3_ModBan | CAAGCAGAAGACGGCATACGATTGATGGTGCCTACAG |
| mRACE for simultaneous mapping of transcriptional start and stop sites | |
| RACE-IcaR-1 | TATCATCAAGTGTGTACCGTCAT |
| RACE-IcaR-2 | TCAAAGATGAAGTGATTTCGCTAC |
| Northern blot probes | |
| NB-IcaR-Sense | AGCTATATCATCAAGTGTGTACCGTCATACCCCTTCTCTG |
| 400-Sense | TGTAATTAGATGACAACTTATTCTTTTCAGGGGAAC |
| RP-T7-Sentido-IcaR | TAATACGACTCACTATAGGGTACTTTCTTCCACTGCTCCA |
| IcaR+1 (BamHI) | GGATCCGAAATATTTGTAATTGCA |
| Deletion of <i>icaR</i> 3'-UTR | |
| IcaR-A (EcoRI) | GAATTCTTTTCAGAGAAGGGGTATGACG |
| IcaR-B (Sall) | GTCGACCAAAAATTATTTCTTCAAAAATAT |
| IcaR-C (Sall) | GTCGACTAATTTATTGATAAGCGCCTA |
| IcaR-D (BamHI) | GGATCCAACGACCACAAAACATACACAA |
| IcaR-E | GAGTAGAAGCAGTATATTTGT |
| IcaR-F | AAATAGAGTGAAGACACCC |
| Deletion of <i>dbpA</i> gene | |
| SA1387-A (BamHI) | GGATCCTCTTATCTTCACCAACATAA |
| SA1387-B (HindIII) | AAGCTTGTGTGTTTTATGTTATTAGG |
| SA1387-C (HindIII) | AAGCTTATTTGGCCTCCTTATATGT |
| SA1387-D (BamHI) | GGATCCCTTTGAAAGTGAAGGAAGA |
| SA1387-E | ACAGTAATGTTTACGACTCA |
| SA1387-F | GTAAGTGATAAGCGACATTA |
| mRNA levels quantification by qRT-PCR | |
| icaC2.SG-5 | TGCGTTAGCAAATGGAGACTATTG |
| icaC2.SG-3 | TGAGAAAGCACTAATCATTGAATGC |
| icaR.SG-5 | TTCTTATTTGAGTTTATTTTCGACATCG |
| icaR.SG-3 | CATAGAAATTTGCTATCTCTTTACTTAATGATTG |
| gyrB.SG-5 | ACGTATGAAGGTGGTACGCATG |
| gyrB.SG-3 | ACGCGTTAATGCACGTTTGA |
| Transcriptional fusions for <i>icaR</i> and <i>icaADBC</i> promoters | |
| icaA-pSA14-Fw (PstI) | CTGCAGCGTACATTCTAATATACCT |
| icaA-pSA14-Rv (BamHI) | GGATCCTAAGCCATATGGTAATTGATAG |
| Constitutive expression of <i>icaR</i> and <i>icaRmΔ3'-UTR</i> mRNAs | |
| IcaR+1 (BamHI) | GGATCCGAAATATTTGTAATTGCA |
| IcaR-Term (EcoRI) | GAATTCCCTTTTAAAAAGATGTGGGTA |
| Constitutive expression of <i>icaR</i> 3'-UTR | |
| 400sinIcaR (BamHI) | GGATCCTTTTTGTTACTAGTTTGTAAATAATTAAC |
| IcaR-Term (EcoRI) | GAATTCCCTTTTAAAAAGATGTGGGTA |
| Constitutive expression of 3XFLAG IcaR protein | |
| IcaR+1 (BamHI) | GGATCCGAAATATTTGTAATTGCA |
| IcaR-3XFLAG-DCHA | GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGACTACAAAGATGACGACG |
| IcaR-3XFLAG-IZDA | ATAAAAAGGATAAGATTATTGATAACGCA |
| IcaR-Term (EcoRI) | ATGATCTTTATAATCACCGTCATGGTCTTTGTAGTCCAATTTTATAACCCCCTACT |

Continued on following page

Table 4. Continued.

| Oligonucleotide | Sequence |
|--|---|
| Constitutive expression of <i>icaR</i> mRNA carrying the UCCCCUG motif deletion | |
| DEL-AntiRBS-IZDA-2 | TGACAACTTATTCTTTTACATTACACTTTTATAATATGTTCAA |
| DEL-AntiRBS-DCHA | ATGTAAAAGAATAAGTTGTCATCTAATTACA |
| Constitutive expression of <i>icaR</i> mRNA carrying the UCCCCUG motif substitution | |
| SUBST-AntiRBS-IZDA-2 | TGACAACTTATTCTTTGTCCCCTACATTACACTTTTATAATATGTTCAA |
| SUBST-AntiRBS-DCHA | ATGTAGGGGACAAAAGAATAAGTTGTCATCTAATTACA |
| Constitutive expression of <i>icaR</i> mRNA carrying the compensatory mutation at the SD region | |
| SUBST-RBS-DCHA | CAGGTCCCCTTTATAAAAATTGAAGGATAAG |
| SUBST-RBS-IZDA | CAATTTTATAAAGGGGACCTGAAAATTAATCACACTATGTTAC |
| Synthesis of the substrates for in vitro assays | |
| RNA T7 <i>icaR</i> +1 | <u>TAATACGACTCACTATAGGGGAAATATTTGTAATTGCAACTTAATTT</u> |
| RP 5'UTR <i>icaR</i> | <u>CTCTGAAAATAAGGTTATTGCG</u> |
| RV | |
| RNA T7 3'UTR | <u>TAATACGACTCACTATAGGGCTTTATATTTGTGAATGGTTAAGT</u> |
| ucccc Fw | |
| RP3'UTR ucccc | TTCCATTTTCATCTAATTTATGCGG |
| Rv | |
| Production of Recombinant RNase III | |
| RNase III Fw (NdeI) | GGAA7TCCATATGTCTAAACAAAAGAAAAGTGAG |
| RNase III Rv (BamHI) | CGCGGATCCCTATTTAATTTGTTTAATTGCTTATAG |
| mRACE for mapping of processing sites at the 3'-UTR | |
| RACE-Frag-400-RT | TTCATCTAATTTATGCGGATTTCTCTG |
| <i>icaR</i> -C (Sall) | GTCGACTAATTTATTGATAAGCGCCTA |
| Toeprint | |
| <i>IcaR</i> Forward | GGGAAATATTTGTAATTGCAACTTAATTTTCC |
| <i>IcaR</i> Reverse | TCCATGGATCCAAAAAGCGCCTATGTCATGATTTAC |
| T7 5'UTR | <u>TAATACGACTCACTATAGGGGAAATATTTGTAATTGCCAACTTAATTT</u> |
| <i>IcaR</i> 5' rev | ATCATCAAGTGTGTACCGTCATACCCCTT |
| T7 3'-UTR | <u>TAATACGACTCACTATAGGGCTTTATATTTGTGAATGGTTAAGTTTGTCTTTGAAC</u> |
| <i>IcaR</i> 3' term | AAAAAGCGCCTATGTCATGATTTACCATCA |
| spa | TTTGCAGCAGGTGTTACGCC |

Italic, restriction enzyme site included in the oligonucleotide. Underlined, T7 promoter

RESULTADOS

4. RESULTADOS

4.1. Identification of long 3'-UTRs in the S. aureus transcriptome

Finding the 5' boundaries of mRNAs is a critical step for transcriptional promoter recognition. Thus, in general, bacterial transcriptome analyses have been focused on the identification of mRNA 5' ends while mRNA 3' ends mapping has been mostly disregarded, limiting our knowledge about the molecular features inside this mRNA region. To overcome this limitation, we have examined genome wide the 3' boundaries of the *S. aureus* transcriptome by combining RNA-seq data obtained in a previous study (Lasa *et al.*, 2011) with tiling array hybridization data of four genetically unrelated *S. aureus* strains. The normalized tiling arrays signals and the mapped reads were integrated in a web repository that enables the visualization of the transcriptome information (*Staphylococcus aureus* Transcriptome Browser, <http://staph.unavarra.es/>). We calculated the 3'-UTR length of each mRNA as the distance between the annotated translational stop codon of the ORF and the last position of RNA reads downstream. Only 3'-UTRs that were present in the four strains analysed by tiling were considered in the analysis and, whenever possible, the position of the predicted intrinsic Rho-independent transcriptional terminator (TT) was calculated according to the TransTermHP v2.07 program (Kingsford *et al.*, 2007). As a result, we identified 1055 mRNAs carrying *bona fide* 3'-UTRs, that is, transcripts ending at an intrinsic TT (Fig. 6A, Fig. 6B and Fig. 7A).

Remarkably, 35,4% of these mRNAs contained 3'-UTRs longer than 100 nt (Figure 6A and C). Also, we found that transcription of about one third of the mRNAs carrying *bona fide* 3'-UTRs may continue downstream the predicted TTs, thus generating a long terminating-read-through-dependent 3'-UTR (Fig. 6B and 7B). These 3'-UTRs showed the longest size and usually overlapped with the mRNAs encoded at the opposite DNA strand. This antisense regulation is subsequently followed by RNA degradation induced by the endoribonuclease III (RNase III) (Lasa *et al.*, 2011). Confirming previous findings in *L. monocytogenes* (Toledo-Arana *et al.*, 2009), we found 24 riboswitch-dependent 3'-UTRs, all of them presenting a length higher than 100 nt (Fig. 6B and 7C). In this type of 3'-UTRs, the TT generated when the riboswitch is in an OFF conformation also acts as the TT of the gene encoded upstream of the riboswitch. As a consequence, the 3'-UTR includes the riboswitch sequence. Taken together, these findings indicate that the presence of long 3'-UTRs is very frequent in the *S. aureus* transcriptome and can generate diverse regulatory mechanisms.

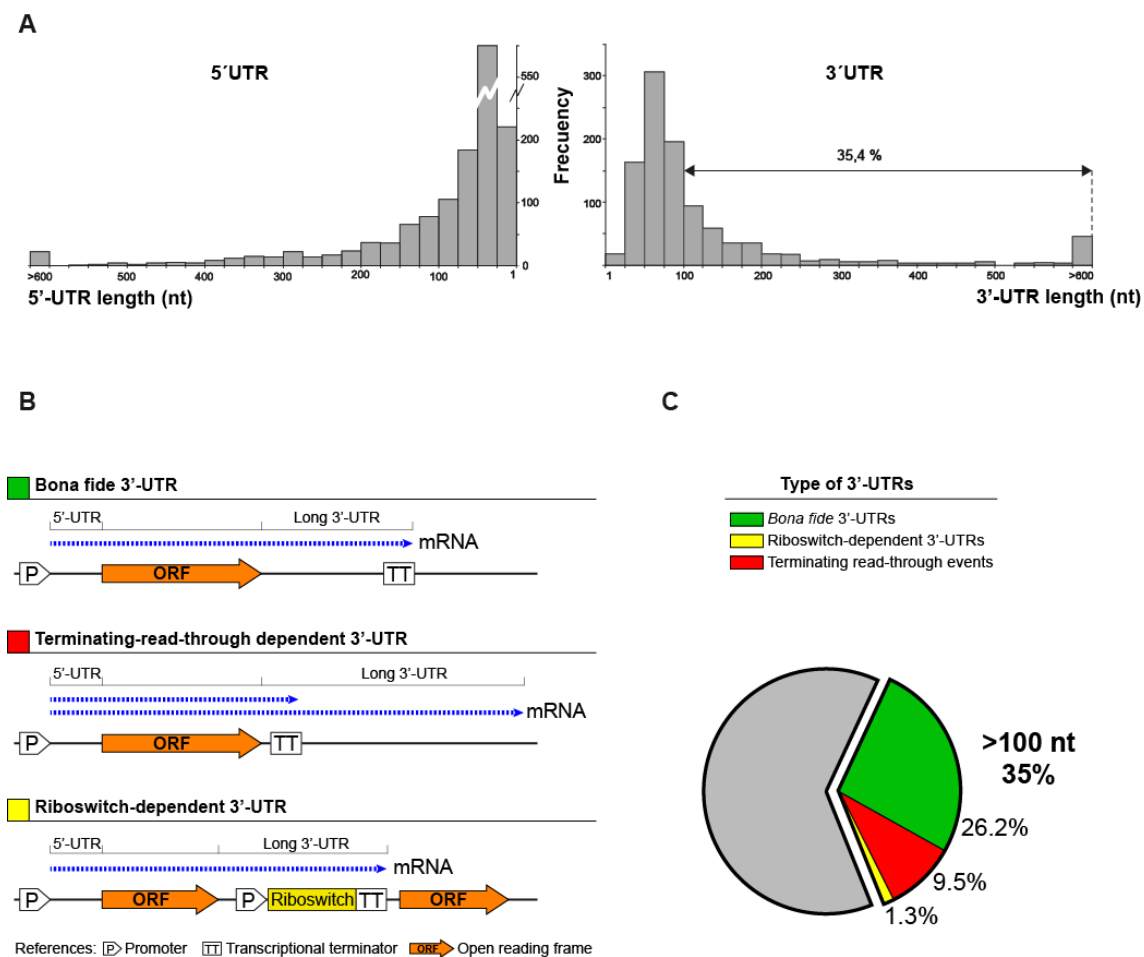


Figure 6. Transcriptome measurements. (A) Histograms of 5' and 3' UTR lengths. Percentage of long 3' UTR (>100 nt) is indicated. (B) Schematic representation of how long bacterial 3' UTRs can be generated. Transcript ending in a transcriptional terminator located far away from the corresponding protein stop codon generates a *bona fide* long 3' UTR. In other cases, despite the presence of a transcriptional terminator (TT) close to the end of the protein stop codon, transcription may continue downstream the predicted TT, generating a terminating-read-through-dependent long 3' UTR. In addition, several transcripts end at a TT that is part of the expression platform of a riboswitch. In this case the long 3' UTRs will be generated only when the riboswitch is in an OFF configuration. Otherwise, if the riboswitch is in an ON configuration, a polycistronic transcript is generated. (C) Pie chart showing types of 3' UTRs.

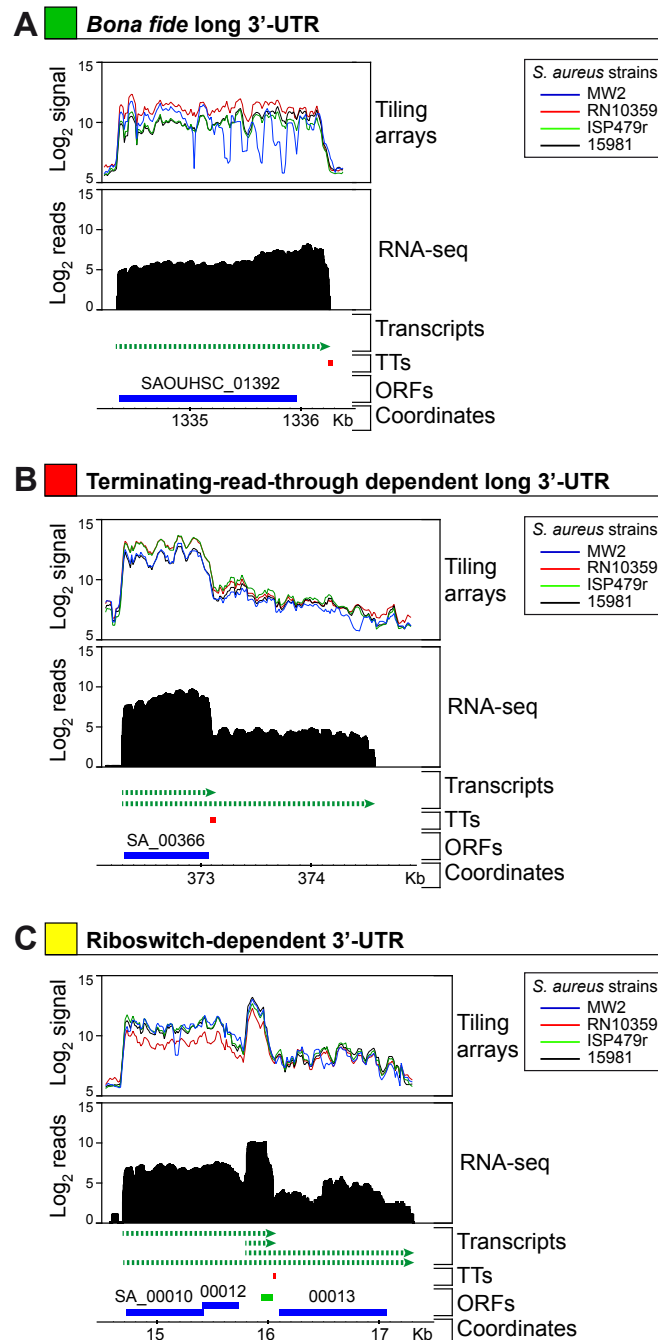


Figure 7. Examples of *S. aureus* transcripts carrying long 3'-UTRs. Drawings are IGB software images showing tiling array signals and RNA-seq mapped reads. The exterior tracks correspond to the tiling signals of four unrelated staphylococcal strains, 15981 (black line), ISP479r (green line), RN10359 (red line) and MW2 (blue line) while the interior tracks (black) correspond to the mapped reads from *S. aureus* 15981 strain. Blue boxes, ORFs; red boxes, intrinsic transcriptional terminators; Green dash lines, transcript. One representative example of a (A) *bona fide* long 3'-UTR; (B) a terminating-read-through dependent long 3'-UTR; (C) a riboswitch-dependent long 3'-UTR.

4.2. 3'-UTRs can be predicted in other bacteria

Our transcriptomic analysis revealed that, in most of the cases, transcription ends in a Rho-independent transcriptional terminator. Therefore, at the end of mapped 3'-UTRs is localized the typical stem-loop RNA structure required to end transcription. These RNA stem-loop structures might be predicted using different programs (Kingsford *et al.*, 2007; Mitra *et al.*, 2010). Knowing the position of these transcriptional terminators relative to the corresponding ORFs, it is possible to determine the putative 3'-UTRs of any sequenced genome. Thus, we calculated the length of the putative 3'-UTRs in different bacterial species as the distance between the stop codon of annotated ORFs and the last nucleotide of the best TT predicted by TransTermHP v2.09 and localized downstream of the corresponding ORF. We considered for the calculations only the TTs predicted with a confidence higher than 95%. Although this value of confidence might sub-estimate the real number of TTs, it decreased the false positive rate of predicted TTs. In any case, if a researcher is interested in a particular gene, this value could be changed or other programs, such as Mfold or WebGesTer (Zuker, 2003; Mitra *et al.*, 2010), can be used to confirm the presence of a TT that TransTermHP algorithm might not detect.

Following this strategy, we analysed 25 genomic sequences representing several prokaryotic phyla. We found that, in 22 out 25 genomes analysed, at least 20% of the predicted 3'-UTRs were longer than 100 nt (Table 5). Remarkably, we could not find a relationship between the 3'-UTR lengths

and the bacterial groups. The content of long 3'-UTRs varies from 9% of *L. monocytogenes* to 63% of *Porphyromonas gingivalis* independently of the phylum, the Gram classification or the % of the AT content. The 3'-UTR lengths are even variable between species belonging to the same genus. For example, in *S. epidermidis* and *S. lugdunensis*, the 54% of predicted 3'-UTRs are longer than 100 nt while in *S. warneri*, *S. carnosus* or *S. pasteurii* only the 22% of them (Fig. 8). The observed differences in 3'-UTR lengths from bacterial species that share many metabolic pathways and orthologous genes suggests that these regions might play roles in bacterial adaptation to different niches. It also indicates that 3'-UTRs might be important for bacterial speciation as it was proposed for eukaryotes (Mazumder *et al.*, 2003; C Y Chen *et al.*, 2012).

Table 5. Identification of 3'-UTRs in different bacterial species based on the prediction of transcriptional terminators.

| Strain | Phylum | Gram | Genome length (nt) | CDS | AT% CDS | AT% IG | TT C>95 | TT L>100 | % Long 3'-UTRs |
|---|----------------|------|--------------------|------|---------|--------|---------|----------|----------------|
| <i>Porphyromonas gingivalis</i> W83 | Bacteroidetes | - | 2343476 | 1909 | 50 | 56 | 327 | 206 | 63,0 |
| <i>Staphylococcus epidermidis</i> ATCC12228 | Firmicutes | + | 2499279 | 2416 | 66 | 72 | 665 | 361 | 54,3 |
| <i>Staphylococcus lugdunensis</i> HKU0901 | Firmicutes | + | 2658366 | 2489 | 65 | 69 | 614 | 326 | 53,1 |
| <i>Clostridium botulinum</i> ATCC19397 | Firmicutes | + | 3863450 | 3550 | 70 | 75 | 814 | 408 | 51,3 |
| <i>Streptomyces coelicolor</i> A3(2) | Actinobacteria | + | 8667507 | 7767 | 27 | 30 | 324 | 137 | 47,3 |
| <i>Staphylococcus saprophyticus</i> 15305 | Firmicutes | + | 2045054 | 2445 | 65 | 70 | 810 | 383 | 47,3 |
| <i>Staphylococcus aureus</i> 8325 | Firmicutes | + | 2821361 | 3055 | 66 | 70 | 590 | 269 | 45,6 |
| <i>Staphylococcus aureus</i> 398 | Firmicutes | + | 2872582 | 2699 | 66 | 70 | 573 | 249 | 43,5 |
| <i>Thermus thermophilus</i> HB8 | Hadobacteria | - | 1849742 | 1908 | 30 | 33 | 100 | 35 | 41,6 |
| <i>Staphylococcus haemolyticus</i> JCSC1435 | Firmicutes | + | 2685015 | 2668 | 66 | 71 | 691 | 271 | 39,2 |
| <i>Streptococcus pneumoniae</i> D39 | Firmicutes | + | 2046115 | 1914 | 59 | 65 | 307 | 118 | 38,4 |
| <i>Cyanobacterium stanieri</i> PCC 7202 | Cyanobacterium | - | 3163381 | 2837 | 59 | 68 | 214 | 48 | 38,3 |
| <i>Chloroflexus aggregans</i> DMS 9485 | Chloroflexi | - | 4684931 | 3731 | 43 | 46 | 318 | 108 | 34,0 |
| <i>Caulobacter crescentus</i> CB15 | Proteobacteria | - | 4016947 | 3737 | 32 | 37 | 374 | 123 | 32,9 |
| <i>Salmonella enteritidis</i> P125109 | Proteobacteria | - | 4685848 | 4206 | 46 | 54 | 613 | 154 | 25,1 |
| <i>Enterococcus faecalis</i> D32 | Firmicutes | + | 2987450 | 2876 | 61 | 66 | 914 | 216 | 23,6 |
| <i>Neisseria meningitidis</i> 8013 | Proteobacteria | - | 2277550 | 1913 | 46 | 54 | 636 | 147 | 23,1 |
| <i>Staphylococcus pasteurii</i> SP1 | Firmicutes | + | 2559946 | 2414 | 66 | 71 | 755 | 173 | 22,9 |
| <i>Lactococcus lactis</i> Il1403 | Firmicutes | + | 2365589 | 2277 | 63 | 69 | 593 | 134 | 22,6 |
| <i>Staphylococcus warneri</i> SG1 | Firmicutes | + | 2486042 | 2335 | 66 | 71 | 760 | 164 | 21,6 |
| <i>Staphylococcus carnosus</i> TM300 | Firmicutes | + | 2566424 | 2461 | 64 | 69 | 835 | 177 | 21,2 |
| <i>Salmonella thyphimurium</i> LT2 | Proteobacteria | - | 4857432 | 4454 | 46 | 54 | 667 | 139 | 20,8 |
| <i>Escherichia Coli</i> K-12 | Proteobacteria | - | 3976195 | 3543 | 47 | 55 | 663 | 129 | 19,5 |
| <i>Bacillus subtilis</i> 168 | Firmicutes | + | 4215606 | 4175 | 55 | 61 | 1340 | 154 | 11,5 |
| <i>Listeria monocytogenes</i> EGD-e | Firmicutes | + | 2944528 | 2867 | 61 | 66 | 884 | 81 | 9,2 |

AT%, percentage of adenines and timines in the coding sequence (CDS) or in the intergenic region (IG). C>95, Number of 3'-UTRs predicted with a confidence higher than 95%. C>95_L>100, Number of 3'-UTRs longer than 100 nt. % Long 3'-UTRs, percentage of predicted 3'-UTRs longer than 100 nt.

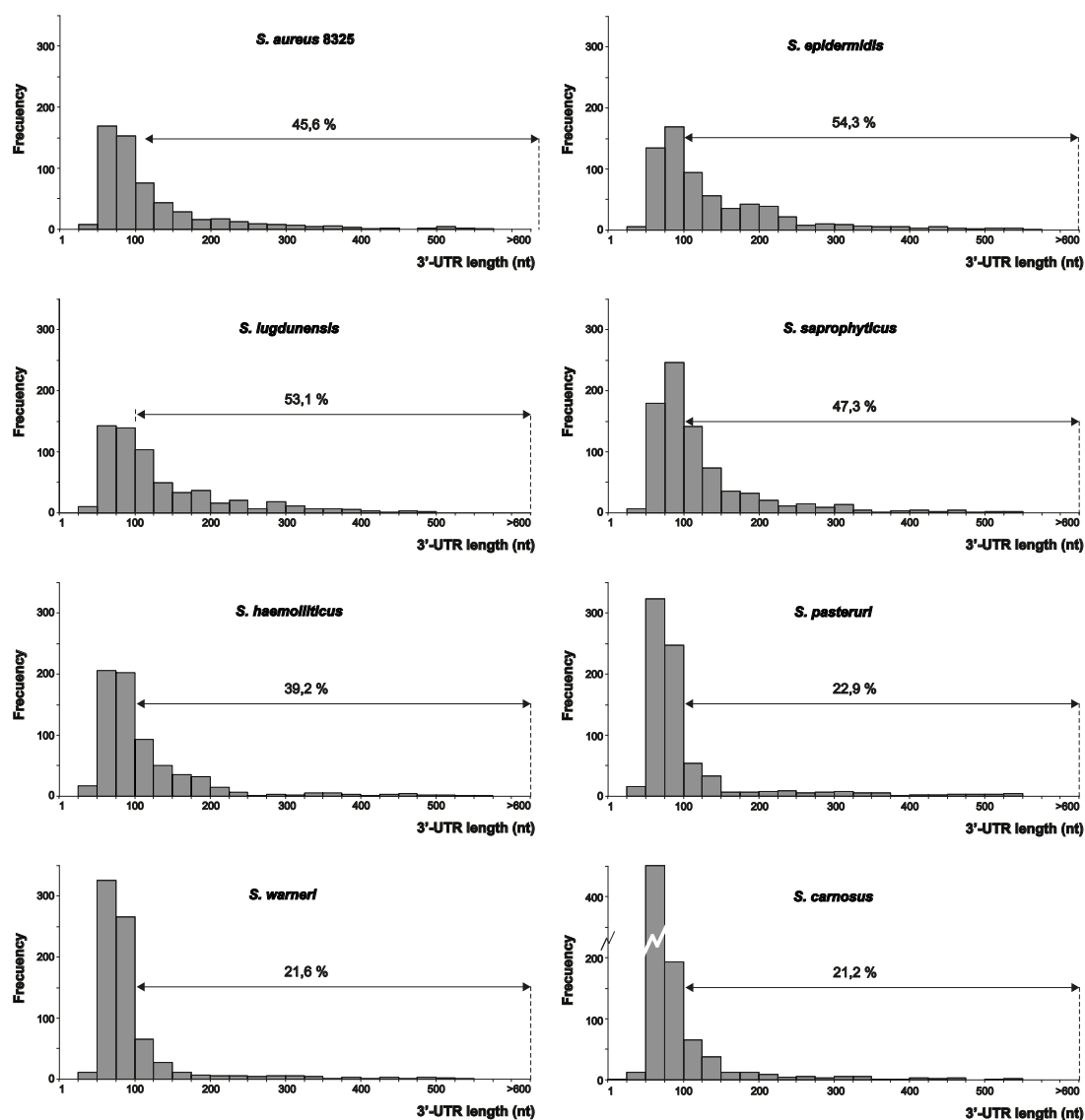


Figure 8 Histograms of predicted 3'UTR length in different *Staphylococci* based in the position of the TTs. All the TTs are predicted with a confidence higher of 95%. The 3'UTR length was calculated as the distance between CDS and the first downstream TT. Percentage of 3'UTR longer than 100 nt is shown.

4.3. The *icaR* mRNA contains a highly conserved long 3'-UTR

To evaluate whether long 3'-UTRs have a regulatory role in *S. aureus*, we focused on *bona fide* long 3'-UTRs and exclusively concentrated on non-overlapping long 3'-UTRs in order to avoid possible effects due to antisense regulation. Due to the relevance of biofilm formation during *S. aureus* infection, we chose the *bona fide* long 3'-UTR of *icaR* transcript (Fig. 9), that encodes for a repressor of biofilm formation, as a model to examine whether long 3'-UTRs play a role in regulating the fate of the mRNA and/or its translation process. We first validated the *icaR* mRNA boundaries using a specialized mRACE protocol that uses circularized RNAs (Toledo-Arana *et al.*, 2009). mRACE experiments located *icaR* mRNA 5' transcriptional start site (TSS) 72-nt upstream from the start codon and the transcriptional termination site (TTS) 390-nt after the stop codon, immediately downstream of the predicted Rho-independent transcriptional terminator (Fig. 10A). Northern blot analysis using either a probe complementary to the 3'-UTR region or to the coding region, confirmed the presence of a band of ~1 Kb that is consistent with the 1023-nt mRNA molecule mapped by mRACE (Fig. 10B). We next excluded that the 3'-UTR codes for a peptide using Glimmer v3.02 at the NCBI web page. From these experiments we concluded that *icaR* mRNA contains a long *bona fide* 3'-UTR that accounts for 38% of the complete mRNA molecule (Fig. 10A).

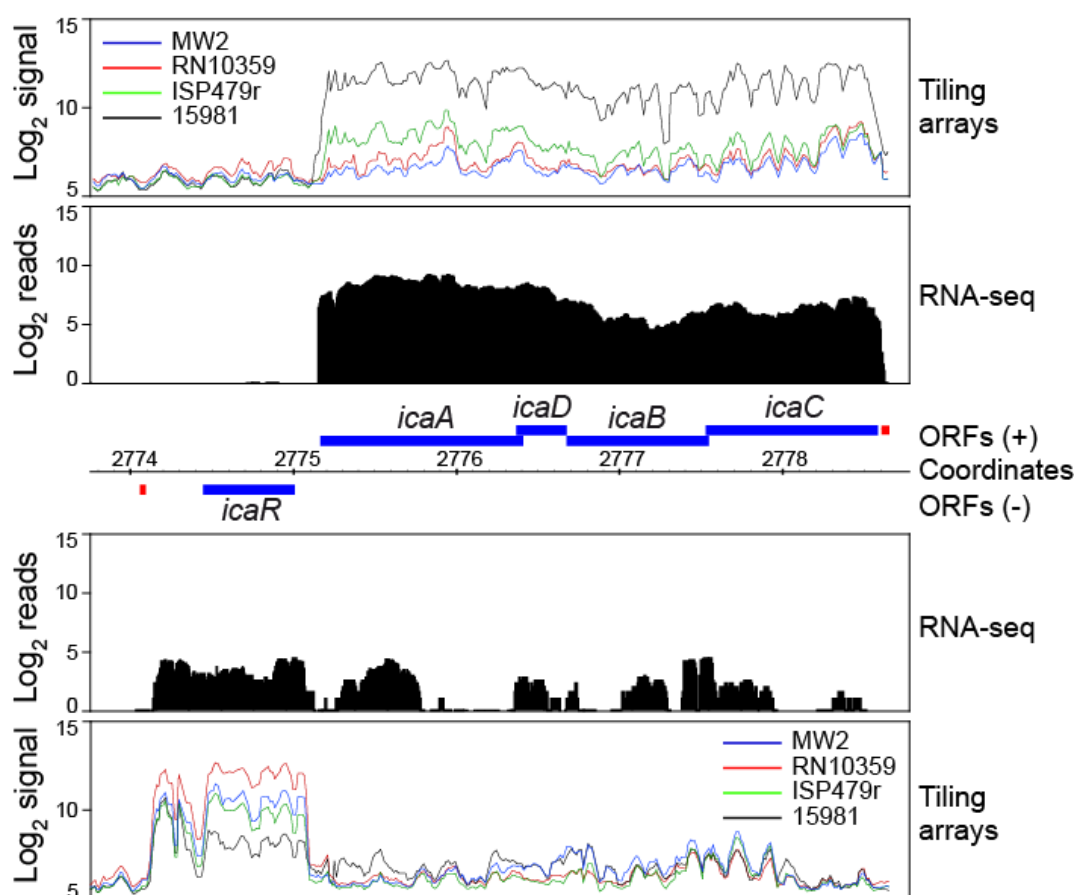
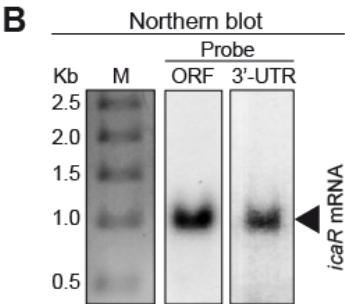


Figure 9. IGB software image showing tiling arrays signals and RNA-seq mapped reads distribution in the *icaRADBC* locus. The exterior tracks correspond to tiling signals of four unrelated staphylococcal strains, 15981 (black line), ISP479r (green line), RN10359 (red line) and MW2 (blue line) while the interior tracks (black) correspond to the mapped reads from *S. aureus* 15981 strain. Blue boxes, ORFs; red boxes, intrinsic transcriptional terminators.



Sequence conservation between intergenic regions of different bacterial species has been traditionally used to identify regulatory non-coding RNAs (Wassarman *et al.*, 2001). Indeed, non-coding regions are more permissive to nucleotide substitutions than protein-coding regions, unless the sequence of the “non-coding” region plays a functional role. We compared sequence conservation of the coding sequence and also of the 3'-UTR inside *icaR* mRNA in the 173 *S. aureus* genomes available at the Microbes database from NCBI. Nucleotide variation analysis showed that 19 nt out of the 561 nt corresponding to *icaR* ORF were variable (3.4%). In the case of the 390 nt 3'-UTR, an accumulation of 26 nt changes occurred, which accounts for a 6.6% of this region. In contrast, the region comprised between the TTS of *icaR* mRNA and the TSS of the neighbour *capA* gene transcript showed a variation of 18.8% (Fig. 11A). A similar analysis was carried out with the RNAIII molecule, which is a multifunctional regulatory RNA that encodes δ -haemolysin, and that also contains a long 3'-UTR of 352 nt acting as an antisense RNA to repress the translation of target mRNAs (Felden *et al.*, 2011). Results again showed a very little nt variation inside both the coding region and the 3'-UTR when compared to the number of changes occurring downstream RNAIII (Fig. 11B). Overall, the high degree of conservation present in the 3'-UTR of *icaR* transcript strengthened our hypothesis that this region might play a functional role in *S. aureus*.

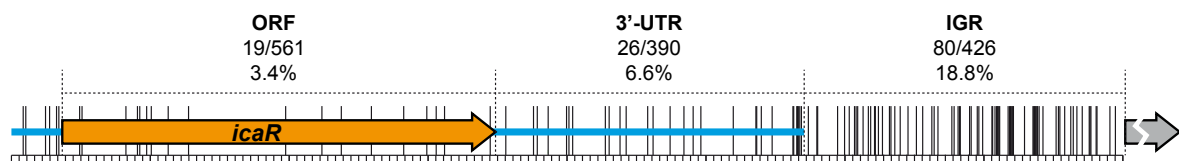
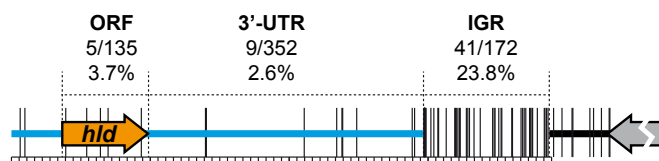
A *icaR* mRNA**B** *hld* mRNA (RNAIII)

Figure 11. *icaR* as well as *hld* 3'-UTRs are highly conserved in *S. aureus*. Nucleotide variation rate was calculated using the 176 *S. aureus* genomic sequences available at NCBI web page. Vertical lines represent a nucleotide change in at least one *S. aureus* genome. **(A)** *icaR* mRNA region. **(B)** *hld* mRNA (RNAIII) region. ORF, open reading frame; IGR, intergenic region comprised between the corresponding 3'-end and the end of the next known transcript.

4.4. 3'-UTR *icaR* post-transcriptionally modulates *IcaR* expression

We then sought to determine whether the long 3'-UTR of *icaR* mRNA modulates the expression of the IcaR protein. We generated an isogenic mutant carrying a 330-bp chromosomal deletion that removes most of the 3'-UTR of *icaR* but preserves the intrinsic TT integrity in the *S. aureus* 15981 strain (Fig. 12A). We then measured *icaR* mRNA levels of wild type and mutant cells grown until exponential phase ($OD_{600nm}=0.8$) by qRT-PCR. Results showed that deletion of the 3'-UTR induced a ~3-fold increase in *icaR* mRNA levels ($P=0.0286$) (Fig. 12B). This increase was also observed by Northern blot (Fig. 12C). Note that the presence of a band corresponding to the expected ~0,7 kb in the Northern-blot implies that the deletion of the 3'-UTR did not affect transcription termination of the mutated *icaR* mRNA (Fig. 12C). Next, we asked whether the increase in the amount of *icaR* mRNA correlated with higher levels of IcaR protein. We tagged the chromosomal copy of *icaR* gene with the 3XFLAG sequence in both the wild type and $\Delta 3'$ -UTR mutant strains. Results showed that $\Delta 3'$ -UTR mutant strain produced significantly higher levels of IcaR protein compared to the wild type strain (Fig. 12D).

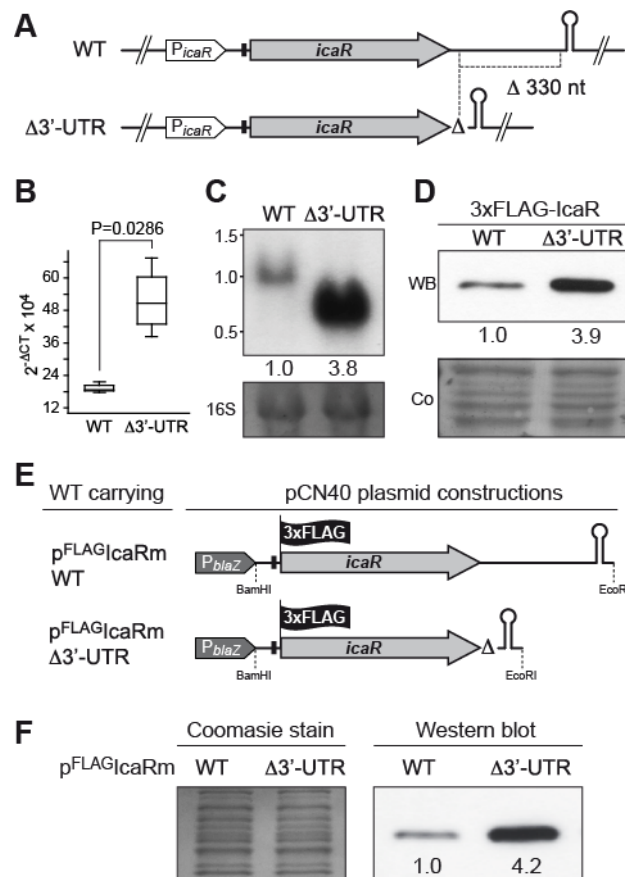


Figure 12. *icaR* 3'-UTR post-transcriptionally modulates IcaR expression. (A) Schematic representation of chromosomal 3'-UTR deletion. Note that the transcriptional terminator is not affected by the deletion. (B) qRT-PCR analysis of *icaR* mRNA levels in *S. aureus* 15981 wild type and $\Delta 3'$ -UTR strains grown in TSB-gluc at 37°C until exponential phase ($OD_{600nm}=0.8$). The *gyrB* transcript was used as an endogenous control, and the results were expressed as the n-fold difference relative to the control gene ($2^{-\Delta Ct}$, where ΔCt represents the difference in threshold cycle between the target and control genes). (C) A representative Northern blot showing *icaR* mRNA of wild type and $\Delta 3'$ -UTR strains grown in TSB-gluc at 37°C until exponential phase ($OD_{600nm}=0.8$). Lower panel shows 16S ribosome band stained with ethidium bromide as loading control. (D) A representative Western blot showing IcaR protein levels expressed from strains shown in panel A. The 3XFLAG tagged IcaR protein was detected with commercial anti-3XFLAG antibodies. Numbers below the image show relative band quantification according to densitometry analysis performed with ImageJ (<http://rsbweb.nih.gov/ij/>). A Coomassie stained gel portion is shown as loading control. (E) Schematic representation of plasmid constructions constitutively expressing the 3XFLAG tagged IcaR protein from the whole *icaR* mRNA or the mRNA carrying the 3'-UTR deletion. (F) A representative Western blot showing IcaR protein levels of strains shown in panel E. The 3XFLAG tagged IcaR protein was detected with commercial anti-3XFLAG antibodies. Densitometry analysis is also shown. On the left, a Coomassie stained gel portion is shown as loading control.

The increase in the amount of *icaR* mRNA/IcaR protein in the absence of the 3'-UTR could be explained either by the existence of higher transcriptional rates from the *icaR* promoter or by an increased stability of *icaR* mRNA or even by a combination of both processes. To distinguish between these possibilities, we uncoupled transcriptional and post-transcriptional regulatory mechanisms by ectopically expressing either the entire *icaR* mRNA or the *icaR* mRNA lacking the 3'-UTR under the control of a constitutive promoter. In both cases, IcaR was tagged with a 3XFLAG epitope at the N-terminal generating plasmids pFLAGIcaRm_WT and pFLAGIcaRm Δ 3'-UTR (Fig. 12E). Western blot analysis revealed that IcaR protein was produced in higher levels (~4 fold) in the strain harbouring pFLAGIcaRm Δ 3'-UTR compared to the strain harbouring pFLAGIcaRm_WT (Fig. 12F). This result suggests that the 3'-UTR mediated regulation occurs at the post-transcriptional level.

We then explored the possibility that the 3'-UTR may reduce mRNA stability. A comparison of *icaR* mRNA stability revealed that *icaR* mRNA half-life increased from 2.1 min in the wild type to more than 10 min in the Δ 3'-UTR strain (Fig. 13A and 13B). Degradation of mRNA can follow several pathways involving a combination of exo- and endoribonucleases and RNA-binding proteins (Belasco, 2010; Arraiano *et al.*, 2010). In order to determine the proteins involved in *icaR* mRNA decay, we compared the relative abundance of IcaR protein in the wild type 15981 *S. aureus* strain and isogenic deletion mutants in proteins that might affect mRNA stability such as PNPase, RNase III, RNA helicases and Hfq.

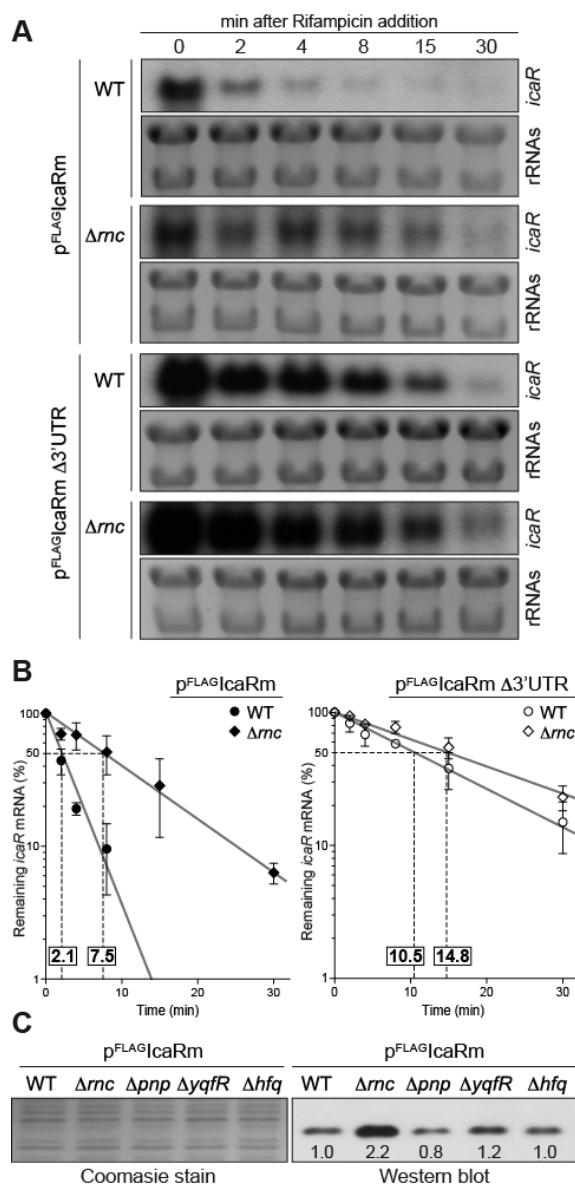


Figure 13. Deletion of *rnc* gene, which encodes the double stranded endoribonuclease RNase III, affects *icaR* mRNA stability and IcaR protein levels. (A) Half-life measurement of *icaR* wild type and $\Delta 3'$ -UTR mRNAs constitutively expressed in the wild type and Δrnc mutant strains. These strains were grown in TSB-gluc at 37°C until exponential phase ($OD_{600nm}=0.8$) and then rifampicin (300 μ g/ml) was added. Samples for RNA extraction were taken at the indicated time points (min). The experiment was repeated three times and representative images are shown. (B) Levels of *icaR* mRNA were quantified by densitometry of Northern blot autoradiographies using ImageJ (<http://rsbweb.nih.gov/ij/>). Each of mRNA levels was relativized to mRNA levels at time 0. The logarithm values of relative mRNA levels were subjected to linear regression analysis and plotted as a function of time. Error bars indicate the standard deviation of mRNA levels from three independent experiments. Dashed lines indicate the time at which 50% of mRNA remained. The half-life of mRNAs is shown above of X-axis. (C) Representative Western blot showing IcaR protein levels in different mutant strains constitutively expressing the 3XFLAG tagged IcaR from the *PblaZ* promoter. Tagged IcaR protein was detected with commercial anti-3XFLAG antibodies. On the left, a Coomassie stained gel portion is shown as loading control. *rnc*, double-stranded endoribonuclease RNase III; *pnp*, polynucleotide phosphorylase PNPase; *yqfR*, (SAOUHSC_01659), ATP-dependent RNA helicase containing a DEAD box domain; *hfq*, RNA chaperone, host factor-1 protein.

Our reasoning was that the IcaR protein should accumulate in the mutants coding for proteins that might be required for *icaR* mRNA decay. The results revealed that only deletion of *rnc*, which encodes RNase III, an endoribonuclease able to degrade double stranded RNA, caused a significant increase in IcaR levels (Fig. 13C). Accordingly, *icaR* mRNA half-life increased from 2.1 min in the wild type to 7.5 min in the *rnc* mutant (Fig. 13A and 13B). Taken together, these results provide compelling evidence that the 3'-UTR is involved in regulating *icaR* mRNA decay in a process that is dependent, at least in part, on RNase III activity.

4.5. Base pairing interaction between 3'- and 5'-UTR of *icaR* mRNA

RNA secondary structures play key roles in post-transcriptional regulatory mechanisms including RNA decay (Arraiano *et al.*, 2010), recruitment of RNA-binding proteins (Babitzke *et al.*, 2009) and riboswitches (Winkler and Breaker, 2005; Serganov and Nudler, 2013). To gain insight into the 3'-UTR mediated *icaR* mRNA decay, the secondary structure of *icaR* mRNA was predicted using the Mfold program (<http://mfold.rutgers.edu/>) (Zuker, 2003). Surprisingly, the prediction showing the lowest energy (initial DG= -230.50 Kcal mol⁻¹) revealed a pairing between the 3'-UTR region (from 890 to 982-nt) and the 5'-UTR SD region (from 4 to 67-nt) of *icaR* mRNA (Fig. 14). The predicted 3'-5'-UTR interacting region comprised around 40 base-paired nucleotides (including G-U interactions).

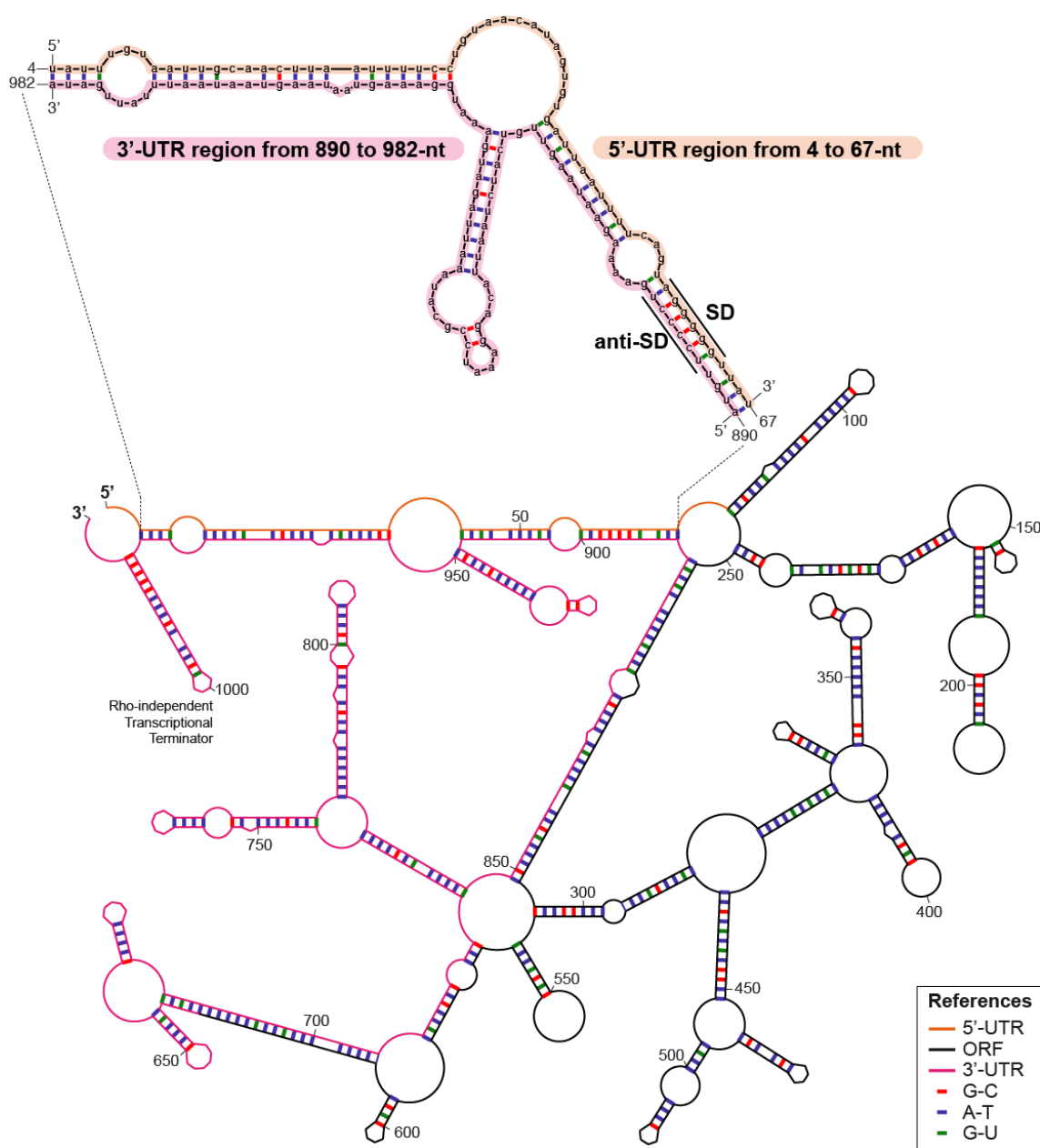


Figure 14. Schematic representation of the RNA secondary structure prediction of the whole *icaR* mRNA molecule generated by Mfold program (<http://mfold.rutgers.edu/>) (Zuker, 2003). G-U pairing was allowed. The 5'-3'-UTRs interaction region is amplified to show pairing nucleotides. The lowest-energy secondary-structural prediction is shown. RNA stem-loops were organized to avoid image superposition.

Because the interaction between the 3'-UTR and the 5'-UTR can be formed either in *cis* or *trans*, we analysed if the full-length *icaR* mRNA was able to fold into various conformations *in vitro*. As expected the mRNA, which was renatured in the TE buffer, migrated as a single band on an agarose gel. In contrast, the *icaR* mRNA which was renatured at 37°C in a buffer containing KCl and MgCl₂, presented two distinct bands (Fig. 15). The slower migrating band corresponded to a molecular mass consistent with a dimeric form of *icaR* mRNA (Fig. 15).

A more detailed analysis of the predicted secondary structure revealed that the pairing region includes a ⁸⁹⁴UCCCCUG⁹⁰⁰ motif located 260-nt downstream of the *IcaR* stop codon complementary to the SD region (⁵⁷UAGGGGG⁶³) (Fig. 14 and 16A). This UCCCC sequence motif has been previously described in several sRNAs of *S. aureus* where it promotes fast binding to target mRNAs and prevents the formation of the ribosomal initiation complex (Boisset *et al.*, 2007; Geissmann *et al.*, 2009). To experimentally monitor the interaction between the ⁵⁷UAGGGGG⁶³ and ⁸⁹⁴UCCCCUG⁹⁰⁰ motifs, we performed RNA gel shift assays using a ³²P-labeled 5'-UTR fragment (117-nt) including the SD sequence and increasing concentrations of a 3'-UTR fragment (120-nt) containing either the wild type ⁸⁹⁴UCCCCUG⁹⁰⁰ or a substituted ⁸⁹⁴AGGGGAC⁹⁰⁰ motif, which disrupts sequence pairing predicted by Mfold program (Fig. 16A and Fig. 16B).

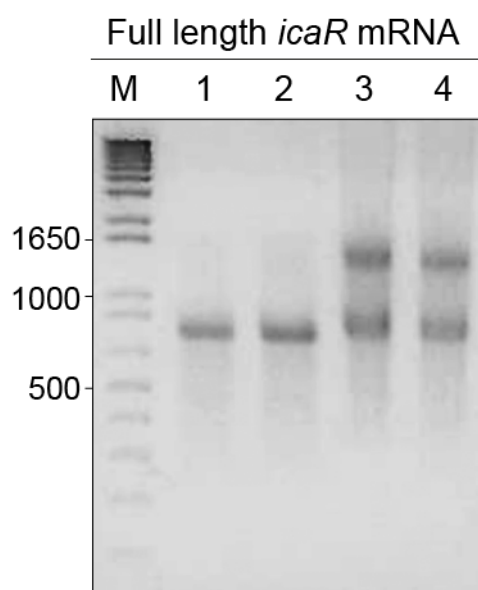


Figure 15. Interaction of two *icaR* mRNA molecules in *trans*. Native agarose gel electrophoresis of the full-length *icaR* mRNA synthesized *in vitro*. **Line 1**, RNA in water; **line 2**, RNA in TE buffer (Tris-HCl 20 mM p 7.5, 1 mM EDTA) incubated 3 min at 90°C, chilled on ice and incubated 15 min at 37°C; **line 3**, RNA renatured in a buffer containing 20 mM Tris HCl pH 7.5, 50 mM KCl at 37°C for 15 min; **line 4**, RNA in a buffer containing 20 mM Tris HCl pH 7.5, 50 mM KCl, 10 mM MgCl₂ renatured at 37°C for 15 min. Sizes of some bands of the molecular weight marker (**M**) are indicated.

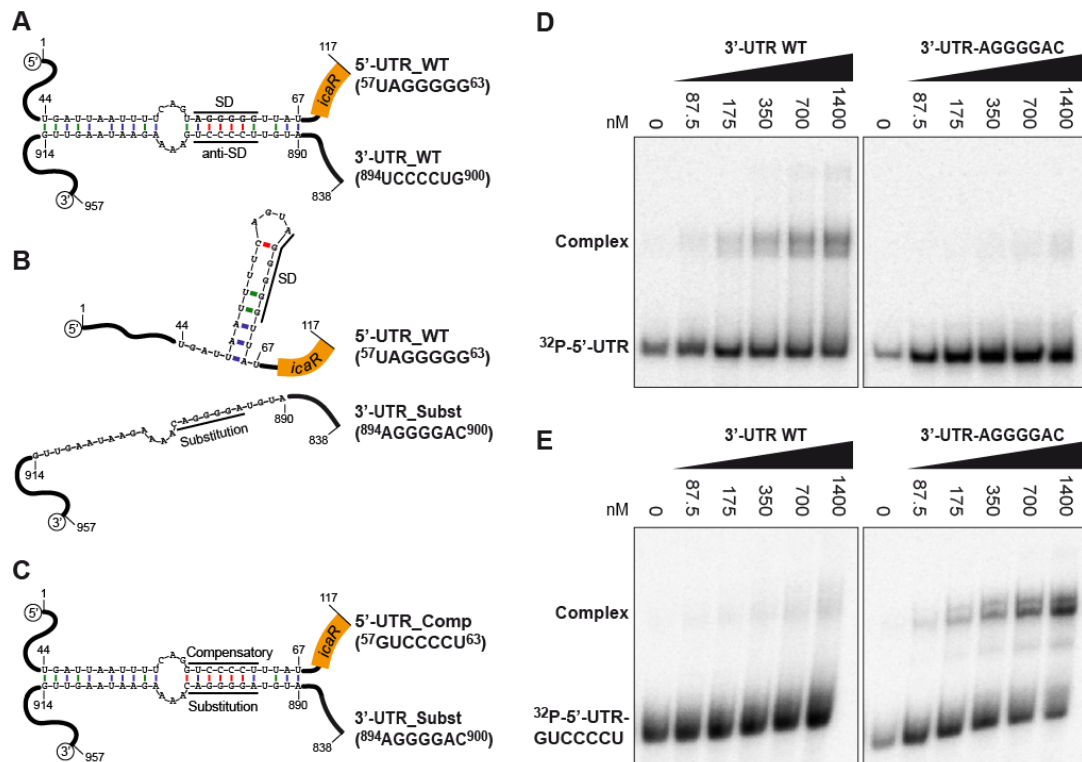


Figure 16. A UCCCCUG motif is necessary for the interaction between the 3'-UTR and the Shine-Dalgarno region of *icaR* mRNA *in vitro*. (A) Schematic representation of the 5'-3' UTRs interaction. (A) UCCCCUG motif located at the 3'-UTR pairs the UAGGGGG Shine-Dalgarno region located at the 5'-UTR. The numbers indicate the relative position of the nucleotides in the full-length *icaR* mRNA (B) Substitution of the ⁸⁹⁴UCCCCUG⁹⁰⁰ motif by ⁸⁹⁴AGGGGAC⁹⁰⁰, disrupts the pairing predicted by Mfold program. (C) Introduction of a compensatory mutation sequence (⁵⁷GUCCCCU⁶³) in the 5'-UTR, complementary to the substituted ⁸⁹⁴AGGGGAC⁹⁰⁰ motif, restores complex formation. (D) Gel shift analysis of the 5' and 3'-UTR *icaR* mRNA interaction. The ³²P-labeled 5'-UTR fragment (1-117-nt) was incubated with increasing concentrations of unlabelled 3'-UTR (3'-UTR WT) or substituted 3'-UTR (3'-UTR-AGGGGAC) (838-957-nt). (E) Similarly, the ³²P-labeled compensatory-5'-UTR fragment (³²P-5'-UTR-GUCCCCU) was incubated with increasing concentrations of unlabelled 3'-UTR (3'-UTR WT) or substituted 3'-UTR (3'-UTR-AGGGGAC). After 30 min of incubation at 37°C, the mixture was analysed by electrophoresis in a native 5% polyacrylamide gel and PhosphorImager (Molecular Dynamics).

The data showed that the 3'-UTR fragment binds to the 5'-UTR fragment carrying the SD with rather low affinity binding (>700 nM). However this interaction is specific because substitution of the ⁸⁹⁴UCCCCUG⁹⁰⁰ motif by ⁸⁹⁴AGGGGAC⁹⁰⁰ severely decreased binding with the 5'-UTR fragment (Fig. 16D). Introduction of a compensatory mutation (⁵⁷GUCCCCU⁶³) in the SD sequence complementary to the substituted ⁸⁹⁴AGGGGAC⁹⁰⁰ motif restored complex formation (Fig. 16C and 16E). Together, these data show that the *icaR* mRNA 3'-UTR specifically anneals to the 5'-UTR in a region that overlaps with the ribosome-binding site.

4.6. The 3'-5'-UTRs pairing provides a substrate for RNase III cleavage

Because pairing between the 3' and 5'-UTRs creates a double stranded region, we sought to determine whether RNase III was capable of cleaving *icaR* mRNA at the pairing region. We performed *in vitro* cleavage assays using a uniformly ³²P-labelled 5'-UTR fragment mixed with the 3'-UTR fragment in the presence of the purified recombinant *S. aureus* RNase III. Results showed that RNase III is able to cleave the 5'-UTR fragment, only in the presence of the 3'-UTR, generating two bands. Thus, we could not detect processed bands either in the absence or in the presence of a 3'-UTR with the AGGGGAC substituted motif (Fig. 17A).

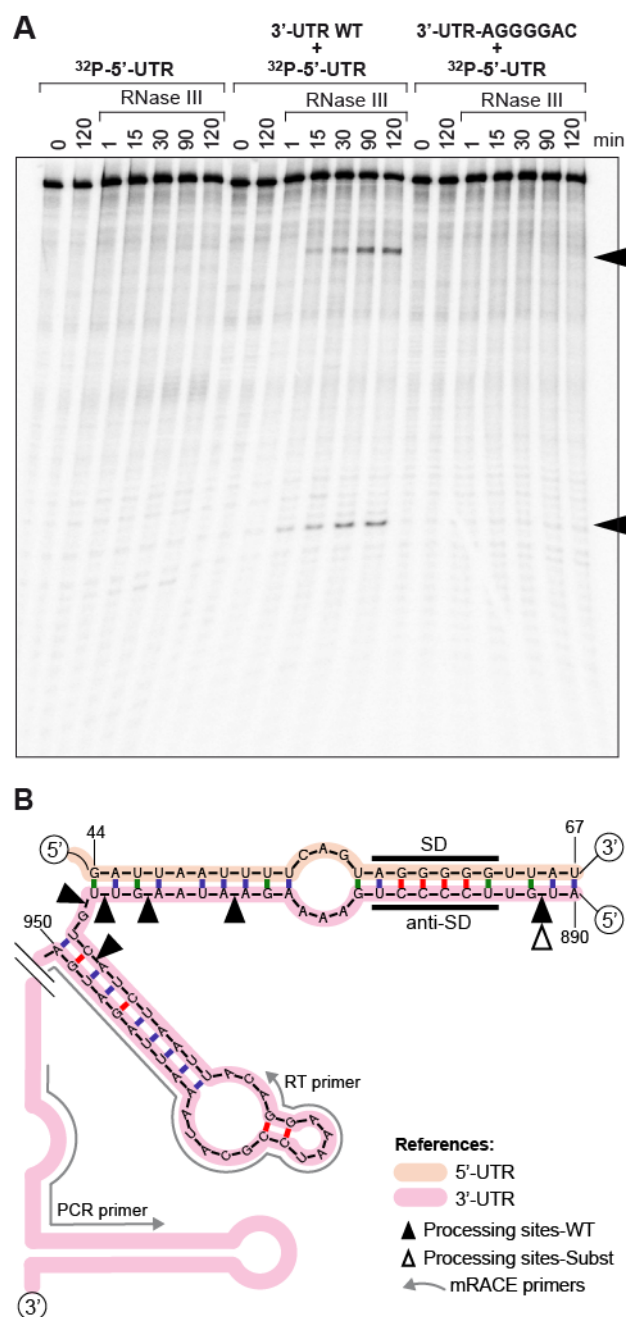


Figure 17. *In vitro* and *in vivo* RNase III-mediated processing of the double stranded region generated by *icaR* 5'-3'-UTRs interaction. (A) *In vitro* RNase III activity assay. A ³²P-labelled 5'-UTR fragment was incubated with purified recombinant *S. aureus* RNase III during different times in the absence or presence of either the 3'-UTR fragment or the substituted 3'-UTR fragment. The two RNA bands that are generated by the presence of the wild type 3'-UTR are indicated with arrows. (B) Schematic representation showing *in vivo* mRACE results. Mapping of *icaR* mRNA fragments naturally generated *in vivo* was carried out with circularized RNAs and two outward primers (RT and PCR) that pair next to the transcriptional terminator. Black and white triangles indicate *in vivo* processing sites identified in the *icaR* mRNA wild type and the *icaR* mRNA with the UCCCC substitution respectively.

To test whether our *in vitro* assay mimicked RNase III capacity to cleave the 3'-5'-UTR duplex of *icaR* mRNA *in vivo*, we performed mRACE analysis with circularized RNA from extracts purified from the wild type, the 3'-UTR *icaR*-SUBST and *rnc* mutant strains. Results showed several processing sites located at the double stranded region of the UCCCC motif when wild type RNA was used (Fig. 17B). In contrast, we were able to detect only one processing site at the UCCCC region in the 3'-UTR *icaR*-SUBST strain. In agreement with the *in vitro* results, no processing sites could be detected in the assay performed with the RNA extract purified from the Δrnc mutant. These results are consistent with the conclusion that RNase III directs the processing of a double stranded region formed by the pairing between *icaR* 3'-UTR and 5'-UTR both *in vitro* and *in vivo*.

4.7. Interaction between icaR 3'-UTR and SD regions prevents the formation of the translational initiation complex

Because the interaction of the 3'-UTR with the 5'-UTR of *icaR* mRNA coincides with the ribosome binding site (RBS), we expected that the 3'-UTR should prevent ribosome loading on the *icaR* mRNA. Toeprint assays were performed to analyse the formation of the ternary ribosomal initiation complexes including purified *S. aureus* 30S ribosomes, initiator tRNA^{Met} and various fragments of *icaR* mRNA. The experiment was first done on a truncated version of *icaR* mRNA containing the whole 5'-UTR and 75 nts from the coding sequence (Fig. 18A).

As expected, the formation of the ternary complex was able to block the elongation of a cDNA primer by reverse transcriptase (RT) to produce a toeprint signal at 16 nt downstream of the initiation codon (Fig. 18A). The addition of increasing concentrations of the 3'-UTR significantly reduced ribosome loading onto the *icaR* 5'-UTR in a concentration-dependent manner. In contrast, increasing amounts of the mutated 3'-UTR *icaR*-SUBST, that cannot form a complex with the *icaR* 5'-UTR (Fig. 16B and Fig. 16D), did not prevent ribosome loading onto the mRNA (Fig. 18A and 18B). We then compared the ability of the *S. aureus* 30S to recognize the 5'-UTR and the whole *icaR* mRNA (Fig. 18C). Quantification of the data showed that the 5'-UTR fragment of *icaR* is recognized by the 30S more efficiently than the full-length mRNA (Fig. 18D). In addition, a RT pause at the SD sequence was slightly stronger with the full-length mRNA than with the 5'-UTR (Fig. 18C). We also have performed toeprinting assays on *S. aureus spa* mRNA which carries a short 5'-UTR, an unstructured ribosome binding site, and a similar SD and initiation codon as found in *icaR* mRNA (Fig. 19A). The data showed that a large proportion of *spa* mRNA was able to form an active initiation complex, and that the 30S recognized *spa* mRNA better than the 5'-UTR and the full-length *icaR* mRNA (Fig. 19B and C). All in all, these results indicate that pairing between the 3'-UTR and the SD region specifically hinders ribosome binding to the *icaR* transcript and that the 5'-UTR is weakly recognized by *S. aureus* 30S.

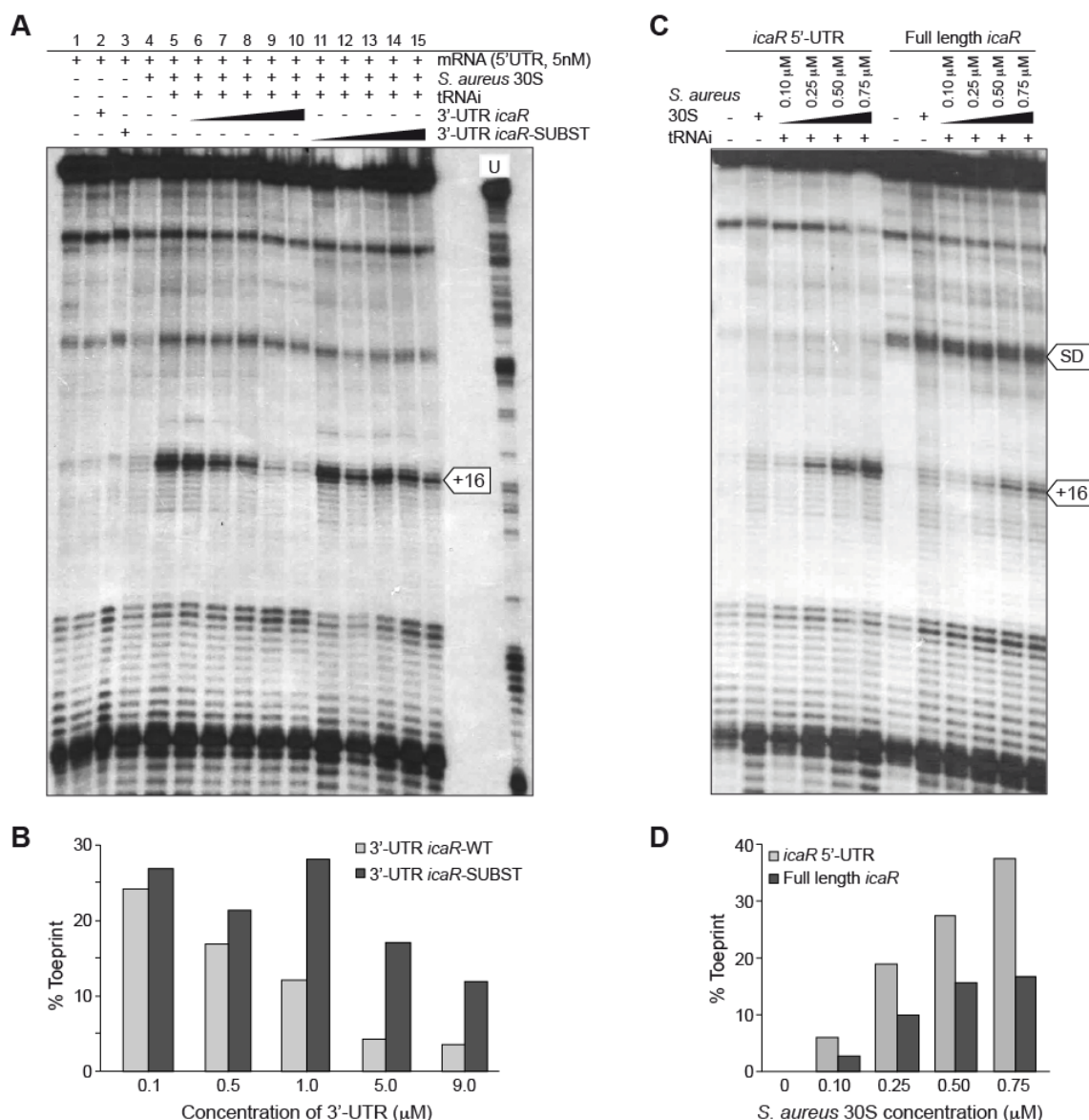


Figure 18. The 5'-3'-UTRs interaction interferes with the translational initiation complex. (A, B) Formation of the ternary complex between *icaR* 5'-UTR fragment (5 nM), *S. aureus* 30S ribosomal subunit, and initiator tRNA was monitored in the absence or in the presence of increasing concentrations of wild-type 3'-UTR fragment and substituted 3'-UTR fragment. The toeprint at position +16 is indicated. The quantification of the toeprint (**B**) was first normalized according to the full-length extension product bands using the SAFA software (Laederach *et al.*, 2007), and the toeprint signal (given in %) represents the yield of the toeprint obtained in the presence of the competitor RNA versus the yield of the toeprint obtained in the absence of the competitor RNA. (**C, D**) Formation of the ternary complex with the 5'-UTR fragment and the full-length *icaR* mRNA molecule was monitored using different *S. aureus* 30S concentrations. A reverse transcriptase pause at the Shine-Dalgarno (SD) sequence occurring in the full-length *icaR* mRNA molecule is indicated with an arrow. The quantification of toeprint experiment (**D**) is described above in B.

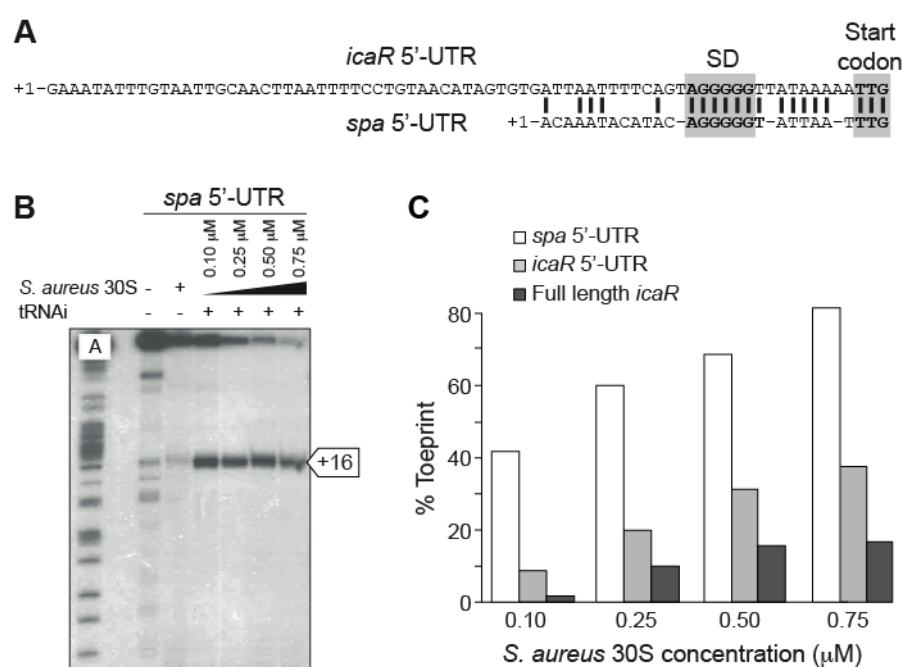


Figure 19. Toeprinting assays performed with *S. aureus spa* mRNA. (A) Sequence of the 5'-UTR of *spa* and *icaR* mRNA are given for comparison. The Shine-Dalgarno (SD) sequence is in bold letter as well as the initiation codon UUG. **(B)** Autoradiography showing the toeprinting assay performed on *spa* mRNA. The toeprint at position +16, representing the formation of ternary initiation complex formed by *S. aureus* 30S ribosomal subunit, mRNA and the initiator tRNA^{Met}, is indicated. **(C)** Comparison of the toeprint signals from panel B and Figure 18C. Band intensity was first normalized according to the full-length extension product bands using the SAFA software (Laederach *et al.*, 2007), and the toeprint signal is given in %.

To demonstrate that the ⁸⁹⁴UCCCCUG⁹⁰⁰ motif is able to regulate IcaR synthesis *in vivo*, we evaluated IcaR protein levels of wild type and RNase III mutant strains harbouring plasmids that constitutively expressed *icaR* mRNA derivatives with a deletion of the ⁸⁹⁴UCCCCUG⁹⁰⁰ motif (p^{FLAG}IcaRm Δ anti-SD) or carrying the substitution of this motif by ⁸⁹⁴AGGGGAC⁹⁰⁰ (p^{FLAG}IcaRm SUBST) (Fig. 20A). Strains expressing *icaR* mRNA derivatives with a deletion or substitution of the UCCCCUG motif accumulated higher levels of IcaR protein, compared to the strain producing wild type *icaR* transcript (Fig. 20B). Interestingly, accumulation of IcaR protein was the highest in the Δrnc mutant strain expressing the *icaR* mRNA with the UCCCCUG substitution (Fig. 20C). These results strongly suggest that both processes, inhibition of ribosome loading and cleavage by RNase III, regulate the IcaR levels.

4.8. Biological relevance of *icaR* mRNA 3'-UTR regulation

IcaR represents the checkpoint of *S. aureus* biofilm formation since it binds to a 42-bp region located just upstream of the *icaA* gene to directly inhibit *icaADBC* transcription (Jefferson *et al.*, 2003). To examine the biological relevance of the *icaR* mRNA 3'-UTR in regulating *in vivo* multicellular behaviour, we first checked if the high IcaR levels observed in the Δ 3'-UTR mutant strain were able to affect *icaADBC* operon transcription.

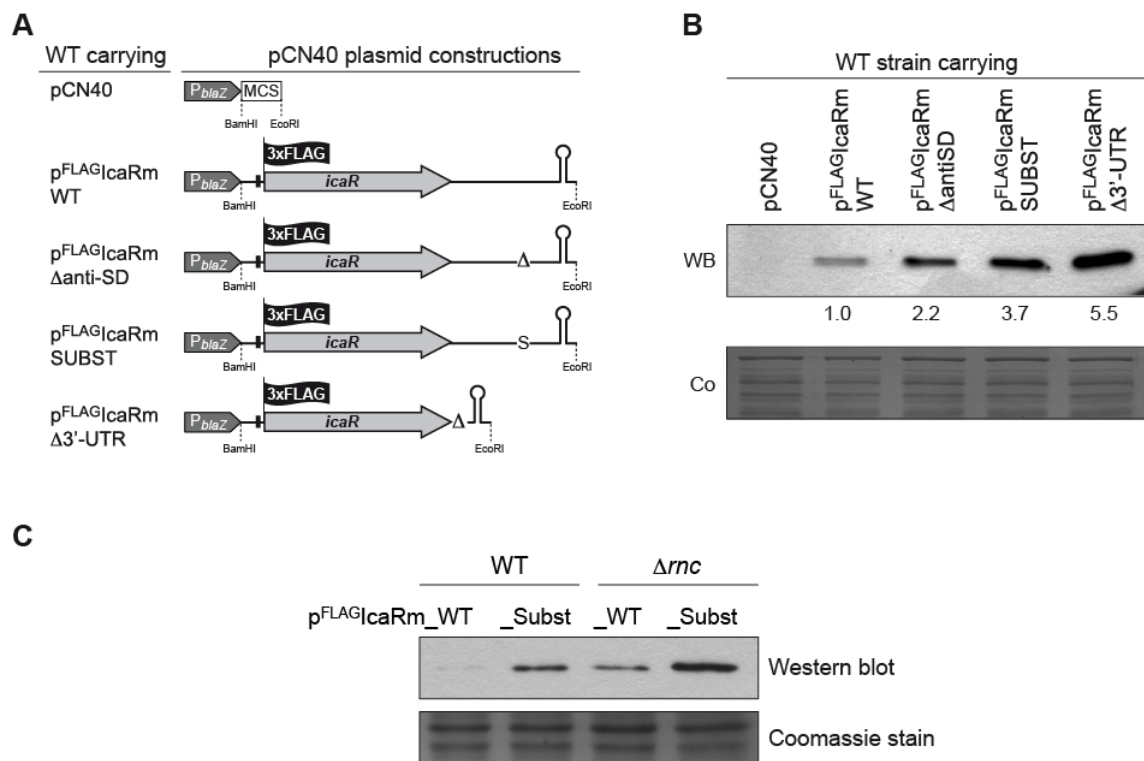


Figure 20. IcaR protein levels expressed from wild type and UCCCC substituted mRNAs in wild type and *rnc* mutant strains. (A) Schematic representation of plasmid constructions constitutively expressing the 3XFLAG tagged IcaR protein from the different *icaR* mRNA alleles. (B) A representative Western blot showing IcaR protein levels in strains shown in panel A. The 3XFLAG tagged IcaR protein was detected with commercial anti-3XFLAG antibodies. Band quantification according to densitometry analysis is shown. A Coomassie stained gel portion is shown as loading control. (C) representative Western blot showing IcaR protein levels in wild type and *rnc* deleted strains constitutively expressing the *icaR* mRNA wild type or the mRNA carrying the UCCCCUG substitution. The 3XFLAG tagged IcaR protein was detected with commercial anti-3XFLAG antibodies. A Coomassie stained gel portion is shown as loading control.

For that, we compared *icaADBC* promoter activity in the wild type strain and its corresponding $\Delta 3'$ -UTR mutant using a transcriptional reporter plasmid comprising the *lacZ* gene fused to the *ica* promoter (*Pica*). As expected, β -galactosidase assays revealed that the activity of the *Pica* promoter was ~7-fold lower in the $\Delta 3'$ -UTR mutant than in the wild type strain (Fig. 21A). Then, we determined the effect of the 3'-UTR deletion on the capacity of two genetically unrelated *S. aureus* strains (*S. aureus* 15981 and *S. aureus* 132) to synthesize PIA-PNAG exopolysaccharide and develop a biofilm. Dot-blot assays using anti PIA-PNAG specific antibodies showed that the synthesis of PIA-PNAG was completely inhibited in both $\Delta 3'$ -UTR mutant strains (Fig. 21B). Accordingly, $\Delta 3'$ -UTR mutant strains lost the capacity to develop a biofilm under continuous-flow conditions in microfermenters (Fig. 21D). Then, to assess specifically the relevance of ⁸⁹⁴UCCCCUG⁹⁰⁰ motif for the regulation of PIA-PNAG synthesis and biofilm development, we tested PIA-PNAG levels and biofilm formation capacity in *S. aureus* 15981 wild type strain transformed with plasmids carrying either full length *icaR* mRNA (plcaRm_WT) or derivatives with deletion or substitution of the UCCCCUG motif (plcaRm_ Δ anti-SD and plcaRm_SUBST respectively). Strains producing *icaR* mRNA derivatives accumulated lower levels of PIA/PNAG (Fig. 21C), and displayed a significant reduction in the capacity to produce biofilm compared to the strain expressing full length *icaR* mRNA (Fig. 21E).

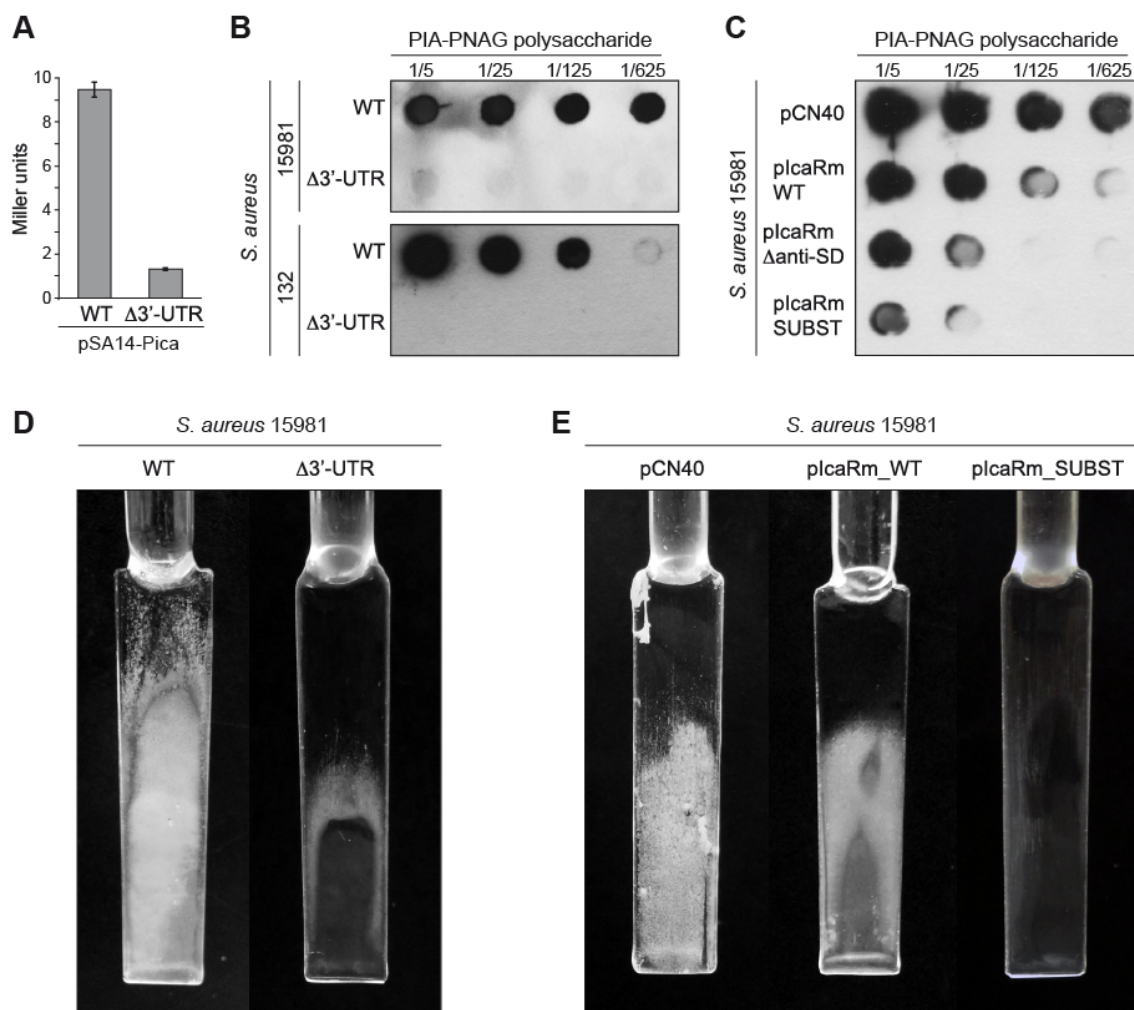


Figure 21. In vivo relevance of the interaction between the 3'-UTR and the Shine-Dalgarno region of *icaR* mRNA. (A) β -galactosidase assays measuring *icaA* promoter activity in the wild type and $\Delta 3'$ -UTR strains grown in TSB-gluc at 37°C until exponential phase (OD_{600nm}=0.8). Each bar represents the average of three independent assays. (B) Consequences of *icaR* mRNA 3'-UTR deletion on PIA-PNAG exopolysaccharide synthesis and biofilm production of *S. aureus* 15981 and 132 strains. (C) In vivo effects of either the mutation or the substitution of the 894UCCCCUG900 motif on PIA-PNAG synthesis. Quantification of PIA-PNAG exopolysaccharide biosynthesis by dot-blot. Serial dilutions (1/5) of the samples were spotted onto nitrocellulose membranes and PIA-PNAG production was detected with specific anti-PIA-PNAG antibodies. (D) Biofilm development of the wild type and $\Delta 3'$ -UTR strains grown in microfermentors under continuous flow for 8 h at 37°C. The glass slides where bacteria form the biofilm are shown. (E) Biofilm development of the *S. aureus* 15981 with the pCN40, plcaRm_WT and plcaRm_SUBST plasmids grown in microfermentors under continuous flow for 8 h at 37°C.

These experiments demonstrate the biological relevance of the *icaR* 3'-UTR and the UCCCCUG motif in controlling PIA-PNAG production and biofilm development by adjusting IcaR repressor protein levels through a post-transcriptional mechanism involving the interaction of a 3'-UTR and a 5'-UTR of the same mRNA.

4.9. Temperature shifts modulate IcaR expression and biofilm formation

Previous results showed that *icaR* mRNA should fold into a secondary RNA structure to promote RNase III dependent processing and to block ribosome loading, preventing the translation of IcaR repressor. It has been shown that secondary RNA structures might suffer conformational changes by external factors, such as temperature shifts, metabolite concentration or the presence of RNA-binding proteins. These conformational changes control the organization of certain transcript regions that include regulatory RNA elements (Noller, 2005; Waters and Storz, 2009; Westhof *et al.*, 2011; Lioliou, Romilly, *et al.*, 2012; Kortmann and Narberhaus, 2012; Lai *et al.*, 2013). Since biofilm formation is an expensive cellular process that should be finely regulated, it might be expected that *icaR* transcript might suffer conformational changes in response to environmental signals to control IcaR expression and thus biofilm formation. Although it has been shown that temperature affects biofilm formation in several bacteria, the effect of this

environmental signal on the *Staphylococcus aureus* biofilm is unknown. Thus, we first analysed PIA-PNAG biosynthesis at 23°C (environmental temperature) and 37°C (host temperature). Dot-blot experiments using anti PIA-PNAG antibodies revealed that *S. aureus* grown at 23°C produced lower amounts of PIA-PNAG exopolysaccharide in comparison with bacteria grown at 37°C (Fig. 22A). Western-blot analysis revealed that IcaR protein is higher expressed at 23°C compared to 37°C (Fig. 22B), explaining the reduction in PIA-PNAG biosynthesis observed at this lower temperature. To investigate if temperature also affects *icaR* mRNA levels, we performed Northern blots, using RNA extracted from *S. aureus* grown at both 23°C and 37°C. As Figure 22C shows, *icaR* mRNA steady-state levels are similar in both conditions suggesting that temperature should modulate IcaR expression at a post-transcriptional level. To validate this result, we analysed biofilm formation capacity at 23°C and 37°C of a strain expressing *icaR* mRNA under the control of the *PblaZ* constitutive promoter (Fig. 22D). We found that this strain form less biofilm at 23°C, in comparison to 37°C, indicating a higher IcaR expression at lower temperature. This result also confirms that *icaR* gene is the only factor required to modulate biofilm formation in a temperature dependent manner.

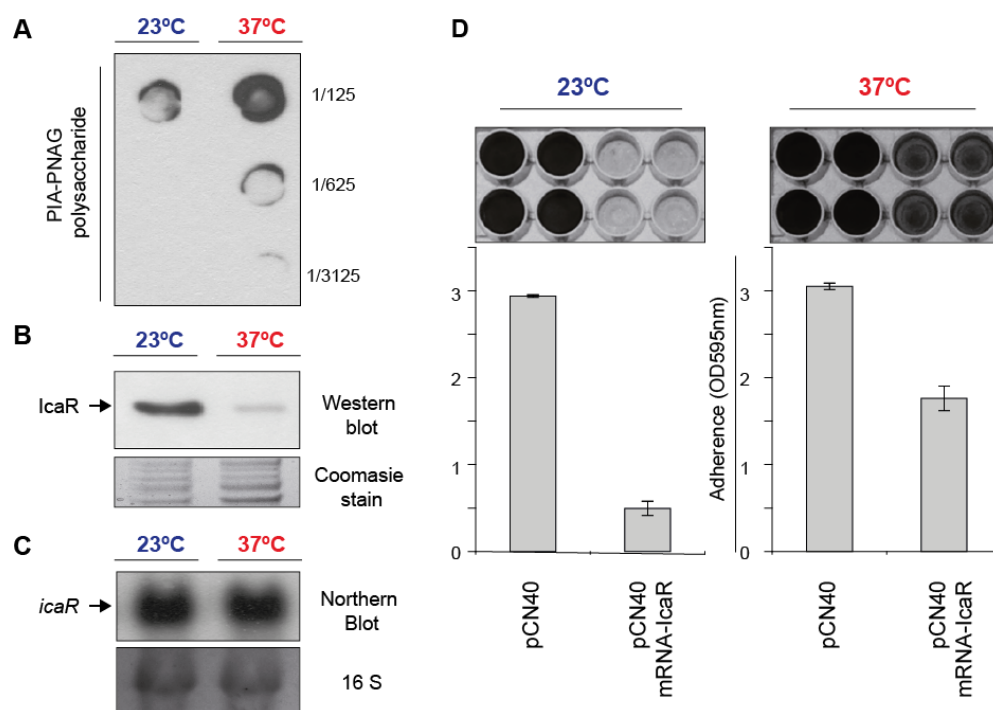


Figure 22. Temperature influence biofilm formation in *S. aureus*. (A) Effect of the temperature on PIA-PNAG exopolysaccharide biosynthesis. Quantification of PIA-PNAG was carried out by dot-blot. Serial dilutions (1/5) of the samples were spotted onto nitrocellulose membranes and PIA-PNAG production was detected with specific anti-PIA-PNAG antibodies. (B) Western blots showing the IcaR protein levels from bacteria grown at 23°C and 37°C. A fraction of a gel stained with Coomassie is shown as loading control. (C) Northern blot assay of steady state *icaR* mRNA levels from bacteria grown at 23°C and 37°C. 16S gel fraction stained with Ethidium bromide is shown as loading control. (D) Biofilm assays on microtiter plates and quantification of violet cristal stain retained by bacteria adhered to the plastic surface.

4.10. Temperature controls the interaction between 5' and 3'-UTRs of *icaR* mRNA

To analyse if temperature might affect the interaction of 3'- and 5'-UTRs, and thus modulate IcaR expression, we performed (in collaboration with Cecilia Arraiano laboratory) RNA gel shifts as described above, but incubating the RNA fragments at 23° and 37°C. The results showed that the 3'-UTR fragment is less efficiently bound to the 5'-UTR fragment at 23°C compared to binding observed at 37°C, indicating that the 5'-3'-UTR interaction is not thermodynamically favoured at 23°C (Fig. 23A). As a consequence, it seems that the double stranded region generated by the 5'-3'-UTR interaction, which is recognized by RNase III, is not formed at 23°C. To test this hypothesis, we performed the RNase III activity assay with RNA fragments that were refolded at 23°C and 37°C. We also included the R1.1 RNA (Amarasinghe *et al.*, 2001) as a control of RNase III activity at lower temperatures. The activity assays carried out with the 5'-UTR fragment in combination with the 3'-UTR wild type revealed that no digestion band could be detect at 23°C as occurred at 37°C (Fig. 23B). The absence of RNase III digestion cannot be explained by a lack of enzymatic activity at 23°C because the pattern of RNase III digestion of the control R1.1 RNA fragment was similar at both temperatures clearly showing that RNase III is active at 23°C (Fig. 23C). Therefore, the absence of RNase III digestion on *icaR* 5'-UTR fragment confirmed that the 5'-3'-double stranded region is not formed at 23°C and might explain the higher levels of IcaR protein observed when bacteria grown at 23°C.

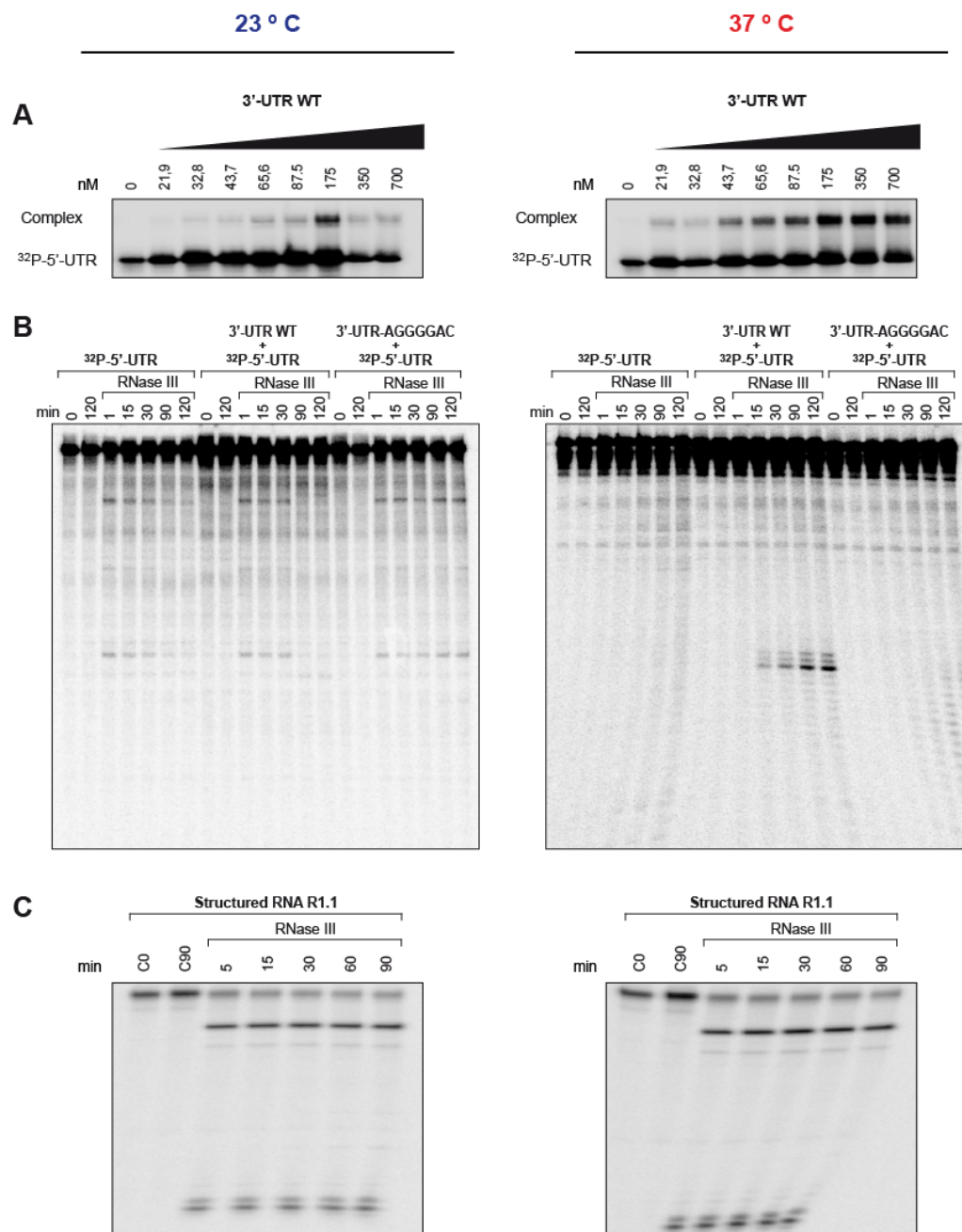


Figure 23. Temperature influences the interaction of 5'-3'-UTRs and subsequently the RNase III cleavage. (A) Gel shift analysis of the 5' and 3'-UTR *icaR* mRNA interaction. The ³²P-labeled 5'-UTR fragment (1-117-nt) was incubated with increasing concentrations of unlabelled 3'-UTR (3'-UTR WT). After 30 min of incubation at 23°C or 37°C, the mixture was analysed by electrophoresis in a native 5% polyacrylamide gel and PhosphorImager (Molecular Dynamics). (B) *In vitro* RNase III activity assay. A ³²P-labelled 5'-UTR fragment was incubated at 23°C and 37°C with purified recombinant *S. aureus* RNase III during different times in the absence or presence of either the 3'-UTR fragment or the substituted 3'-UTR fragment. The two RNA bands that are generated by the presence of the wild type 3'-UTR are indicated with arrows. (C) *In vitro* RNase III activity assay carried out with a standard substrate R1.1 to show that RNase III is active at 23°C.

4.11. *icaR* 5'-UTR forms a three way-helix RNA structure

Considering that temperature influences the pairing of *icaR* 5'-3'-UTRs, and thus the processing mediated by RNase III, we assumed that the conformational changes of *icaR* mRNA structure should have a direct implication on the regulation of IcaR protein translation. Since the 5'-3'-UTRs interaction is not occurring at 23°C, we thought that an alternative structure might be formed at the 5'-UTR, which avoids the interaction with the 3'-UTR. Therefore, in collaboration with the laboratory leaded by Prof. Pascale Romby, the secondary structure of the first 144 nt of the *icaR* mRNA, which includes the whole 5'-UTR (72 nt) and the first 24 codons of the IcaR protein (72 nt from the coding region), was probed at 23°C and 37°C using different enzymes and chemicals under conditions where less than one cleavage/modification per molecule takes place. The Fig. 24A shows the chemical probing experiments carried out on the *icaR* 5'-UTR substrate alone with DMS (which alkylates cytosine at N3 and adenine at N1) while the Fig. 24B shows the enzymatic probing experiment carried out on the same RNA substrate, which was denatured and refolded at 23°C and 37°C before the addition of enzymes RNase T1 (cleaves at unpaired guanine), RNase T2 (cleaves at unpaired adenine) and RNase V1 (cleaves double stranded RNA) respectively. Then, the Figure 24C shows the deduced secondary structure of the *icaR* mRNA 5'-end fragment inferred from the reactivity levels of the enzymatic and chemical compounds used on the probing experiments (Fig. 24 A and B).

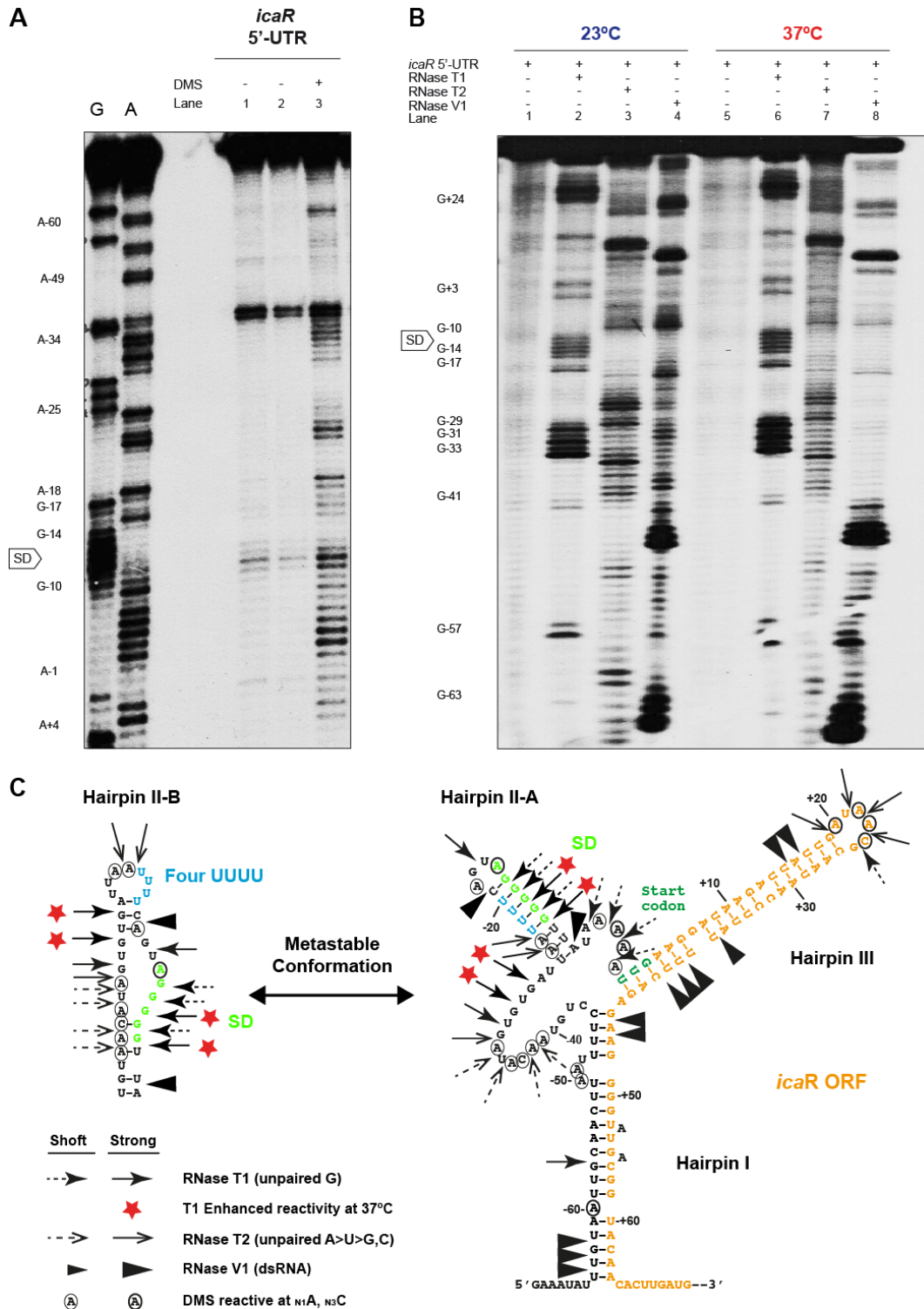


Figure 24. *icaR* 5'-UTR secondary structure. (A) Autoradiogram showing the chemical modifications with DMS obtained by probing *in vitro* the 5'-UTR of *icaR* mRNA. (B) Autoradiogram showing the enzymatic hydrolysis with RNase V1, T1 and T2 revealed by probing *in vitro* the 5'-UTR of *icaR* mRNA. (C) Model of the *icaR* 5'-UTR secondary structure inferred from the probing data. Positions of the cleavages are indicated accordingly. The SD motive (fluorescent green) the start codon (green) and the four U (blue) and the *icaR* ORF (orange) are highlighted.

Based on these assays we could predict that the 5' extreme of the *icaR* mRNA contains several hairpins or stem-loops (Fig. 24C). The DMS probing data was well correlated with the existence of three hairpins adopting a conformation similar to three-way helical junctions (Lescoute and Westhof, 2005; Pena *et al.*, 2009). The nucleotides that are reactive to DMS are mapped on single-stranded mRNA regions corresponding to hairpin junctions, bubbles or apical loops of the *icaR* mRNA fragment. The hairpin I is formed by the pairing of the 5'-end (nucleotides -65 to -44, relative to the start codon) with a region located inside of the *IcaR* ORF (nucleotides +45 to +64). This double-strand region is also confirmed by RNase V1 cleavages at positions -62 to -65, +45 and +46. This hairpin I creates a conformation that enclosed hairpin II and III. The hairpin III comprised the first 14 codons of the *IcaR* ORF (nucleotides +1 to +43). The conformation of this hairpin is also well supported by RNase V1 cleavages found at positions +15, +16, +36 to +40, and by RNase T1 cuts at G+24 and RNase T2 cuts at A+19, A+21, A+22, and C+23, which confirms the accessibility of the apical loop observed by DMS. The conformations observed at the SD region, which correspond to hairpin II, are much more complex. Lower chemical and enzymatic probing signals (DMS reactivities mainly at positions -39 to -34 and RNase T1 cleavages mainly at positions G-14 to -10, respectively) indicated that this secondary structure is rather unstable and might alternate between two conformations. One of them (hairpin IIA, from -26 to -7) creates a stem-loop that blocks the SD region by a FourU motif (4U, positions -20 to -23). Remarkably, this 4U motif has been described in RNA

thermosensor domains usually located at 5'-UTRs of heat-shock and virulence genes (Waldminghaus *et al.*, 2007; Kortmann and Narberhaus, 2012). The second conformation (hairpin IIB, from -42 to -7) is more relaxed being the SD region unpaired. We thought that this metastable hairpin II might act as a thermosensor region refolding the *icaR* three-way helix junction and thus modulating 3'-UTR interaction. However, probing experiments carried out with the *icaR* 5'-end RNA fragment refolded at 23°C revealed that only G-10, G-12, G-29 and G-31 becomes more accessible to RNase T1 cleavage indicating that *icaR* 5'-end region alone is not sufficient to significantly reorganized its structure when a temperature shift occurs.

4.12. The three way-helical junction is disrupted by the binding of 3'-UTR in a temperature dependent manner

It is important to notice that the hairpin II correspond to the 5'-end region of the *icaR* mRNA molecule, which is required to pair with the 3'-UTR (Fig. 14 and Fig. 16). Therefore, changes on this metastable conformation should be important to determine precisely when the 3'-UTR and the ribosome will interact with the SD region. Since temperature shifts does not significantly change the three-way helical conformation of the 5'-end *icaR* mRNA fragment alone, we hypothesized that this structure could be only reorganized by the presence of the 3'-UTR region. In fact, three-way helices fold into well defined structures like ribozymes, riboswitches, RNase P, rRNAs, group II introns or the signal recognition particle (Lescoute and

Westhof, 2005; Klein and Ferré-D'Amaré, 2006; Pena *et al.*, 2009; Lee and Groisman, 2010; Breaker, 2012; Serganov and Nudler, 2013). This RNA conformations have been described to adopt a characteristic Y silhouette in which two hairpins stay immobile to allow the remote interaction of the third flexible helix with other mRNA regions (Lescoute and Westhof, 2005; Pena *et al.*, 2009; Westhof *et al.*, 2011; Butcher and Pyle, 2011; Draper, 2013). Regarding *icaR* mRNA molecule, it could be possible that flexibility of hairpin II is required for interaction with 3'-UTR. Therefore, we performed similar chemical and enzymatic probing experiments of the 144-nt *icaR* 5'-end, this time in presence of the 3'-UTR fragment (Fig. 25A and B). The results showed that, when *icaR* 5'- and 3'-end fragments are refolded at 23°C, both chemical and enzymatic probing patterns of the *icaR* 5'-end are similar in presence or absence of the 3'-UTR fragment (Fig. 25A and B), which confirms that 3'-UTR cannot bind the 5'-UTR at this temperature and thus none conformational changes are induced by the presence of the 3'-UTR (Fig. 26A). In contrast, when both RNA fragments are refolded at 37°C, the 5'-end suffers a drastic structural reorganization in presence of the 3'-UTR, which is revealed by a rearrangement of hairpin I and II that completely open to pair with the 3'-UTR while hairpin III remains stable (Fig 25A, Fig. 25B and Fig. 26B). As a consequence, the binding of the 3'-UTR induces protection of DMS and RNase T1 reactivities around the SD region (hairpin II), mainly at positions A-18 and G-12 to G-14 respectively. At the same time, clear enhanced DMS reactivities were found at positions A-49, A-50, A-53 and A-54 that correspond to the former hairpin I.

Resultados

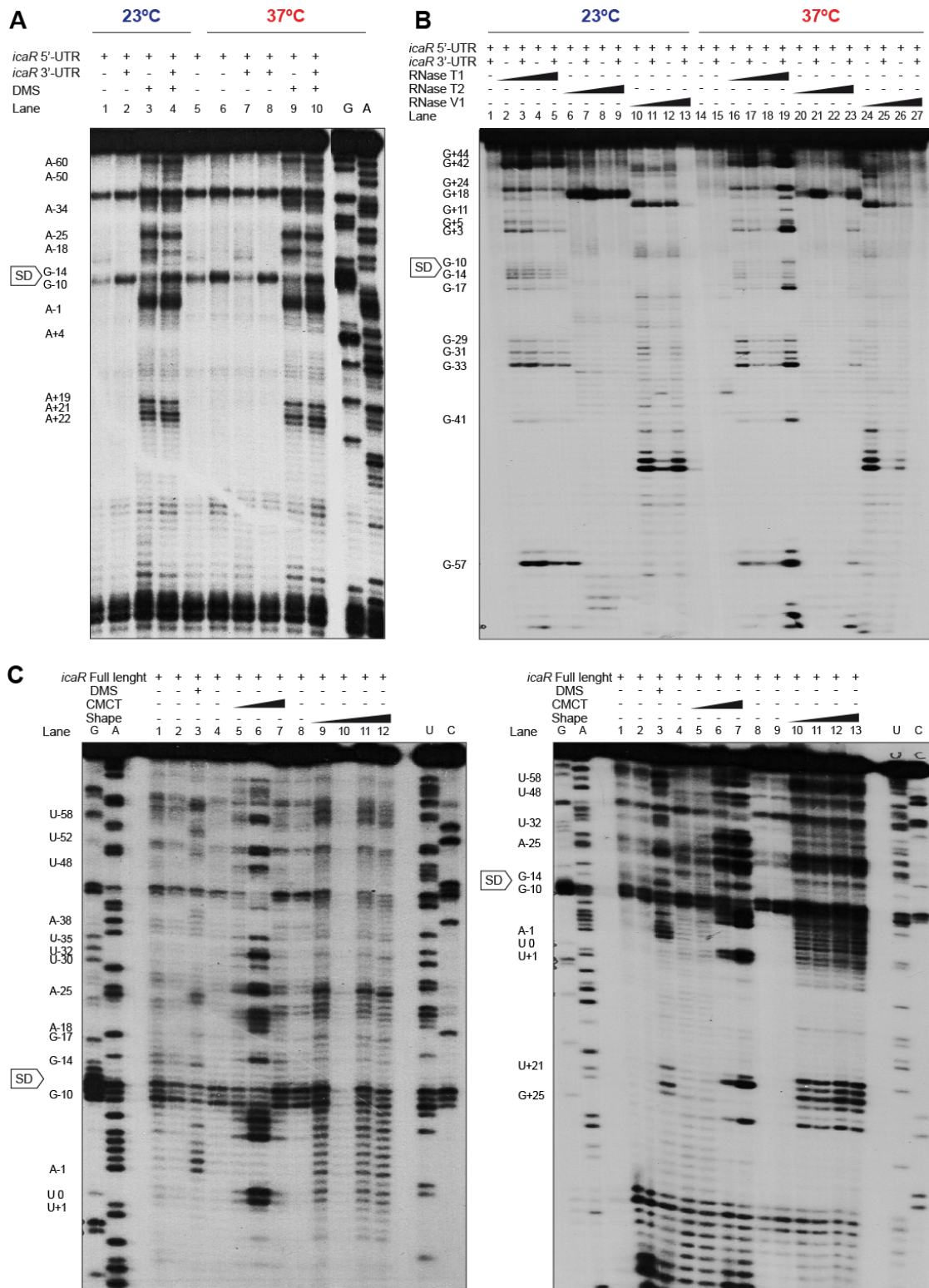


Figure 25. The 5'-end of the *icaR* mRNA suffers a structural rearrangement by the presence of the 3'-UTR at 37°C. (A) DMS probing of the 5'-end region in presence or absence of the 3'-UTR at different temperatures. (B) Enzymatic probing with RNase T1, T2 and V1 of the 5'-end region in presence or absence of the 3'-UTR at different temperatures. (C) The full-length *icaR* mRNA was probed at 37°C with the following chemicals: DMS (N_1A , N_3C), CMCT (N_3U , N_1G) and Shape (unpaired riboses).

In addition, these results could be also validated by probing the 5'-end region of the full-length *icaR* mRNA molecule at 37°C with DMS, CMCT and SHAPE technics (Fig 25C and Fig. 26B). Altogether, these results showed that the AGGGGG SD region might be competitively paired by two motifs, the 4U and the UCCCC, which are separated by 840-nt but both encoded in the same mRNA molecule. The three-way helical conformation adopted by the *icaR* 5'-end region might act as a translational brake uncoupling translation from transcription and giving to transcribe the 3'-UTR. Once the 3'-UTR is transcribed, this region might interact with the three-way helical conformation at 5'-UTR in a temperature dependent manner.

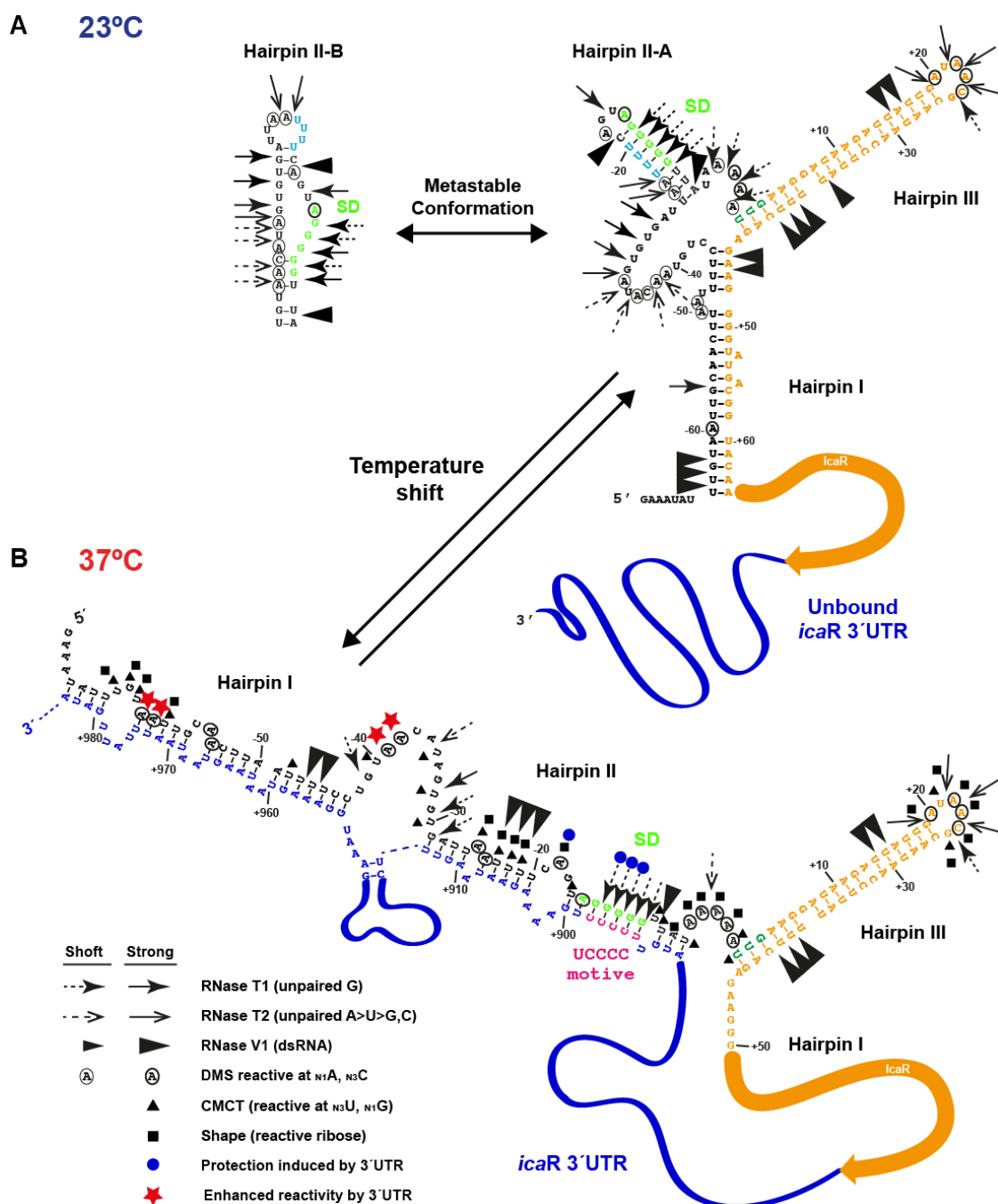


Figure 26. Model of the secondary structure of the *icaR* 5'-end region at 23°C and 37°C inferred from chemical and enzymatic probing. (A) Secondary structure of the *icaR* 5'-end region at 23°C. This conformation does not allow the interaction with the 3'-UTR region but should allow translation of IcaR. (B) Secondary structure of 5'-end of *icaR* paired with the 3'-UTR region at 37°C. Nucleotides that show enhanced reactivity upon the binding of the 3'-UTR are denoted with a red star. In the other hand, nucleotides that has been protected by the 3'-UTR pairing are highlighted with a blue circle. The 5'-UTR (black), the ORF (orange), the 3'-UTR (blue), the SD (fluorescent green), the start codon (green UUG), the FourU (light blue) and the UCCCC (pink) motifs are highlighted.

DISCUSIÓN

5. DISCUSIÓN

5.1. Regulación de la traducción mediante la interacción entre 3'- y 5'-UTRs

Los análisis bioquímicos y genéticos llevados a cabo en este estudio revelan que el proceso de traducción de la proteína IcaR depende de la capacidad de la 3'-UTR larga del mRNA de *icaR* para interactuar con una región complementaria de unos 40 nt presente en la 5'-UTR del mismo mRNA. Así, la inhibición de esta interacción, ya sea mediante la delección o la substitución de unos pocos nucleótidos de la región complementaria, o naturalmente cuando la bacteria crece a 23°C, provoca el aumento de la expresión de la proteína IcaR *in vivo*. Debido a que los motivos que interactúan corresponden a las UTRs que están codificadas en la misma molécula del mRNA, es lógico plantear la pregunta de si la interacción se produce de manera intramolecular (la 3'-UTR con la 5'-UTR de una misma molécula de mRNA) o intermolecularmente (la 3'-UTR de una molécula de mRNA con la 5'-UTR de otro mRNA). Con respecto a la primera posibilidad, es decir la circularización del mRNA, o dicho de otro modo, la formación de un puente físico entre los extremos 5' y 3', es un mecanismo de regulación traduccional ampliamente aceptado en los eucariotas y virus. En ellos, esta interacción se produce generalmente a través de proteínas que se encuentran asociadas tanto con el extremo 5' como con el 3' del mRNA. Dependiendo de las proteínas que están involucradas, la interacción puede

estimular o reprimir el proceso de traducción de la proteína codificada en el mRNA (Mazumder *et al.*, 2003). Alternativamente, la circularización de transcritos que se generan a partir de algunos virus de RNA, puede iniciarse de una manera mucho más sencilla, sin la necesidad de la participación de proteínas accesorias. Por ejemplo, durante el proceso de traducción de virus, como el luteovirus del enanismo amarillo de la cebada o el virus del dengue, el genoma de RNA de cadena positiva forma un círculo cerrado gracias a la interacción directa entre las regiones complementarias localizadas en ambas UTRs. Esta circularización del RNA permite el inicio de la traducción en la región 5' próxima al codón de iniciación (L Guo *et al.*, 2001; Alvarez *et al.*, 2005). Este escenario no se puede comparar de manera directa con la interacción entre las 5'-3'-UTRs del mRNA de *icaR*, porque dicha idea entraría en conflicto con el dogma que establece que la transcripción y la traducción están acoplada en bacterias. Dicho de otro modo, si los ribosomas comienzan la traducción del mRNA de *icaR* antes de que la RNA polimerasa sintetice el motivo UCCCC, la 3'-UTR no sería capaz de modular las rondas iniciales de traducción de IcaR, a no ser que exista un mecanismo que permita desacoplar la transcripción de la traducción del mRNA de *icaR*. Experimentos comparativos que miden el acceso de los ribosomas a la RBS muestran que la región 5'-UTR de *icaR* es reconocida menos eficientemente que la 5'-UTR del mRNA de *spa*, que se caracteriza por tener una RBS no estructurada. Estos datos, sumados a las predicciones de estructura secundaria realizadas mediante análisis químicos y enzimáticos, muestran

que la 5'-UTR de *icaR* puede adoptar una estructura secundaria, que incluye tres horquillas de RNA, que ralentizaría el acceso de la 30S durante el proceso de transcripción, pero que es lo suficientemente inestable como para ser desplazada por la 3'-UTR como muestran los ensayos de retardo en gel. Este mecanismo secuencial sería similar al control temporal de la traducción descrito para el mRNA de *hok*, el cual está involucrado en la muerte celular programada (Møller-Jensen *et al.*, 2001). Por supuesto, tampoco se puede excluir que el inicio de la traducción pudiese estar inhibido por otro mecanismo mientras el mRNA de *icaR* se transcribe completamente. Por ejemplo, la propia estructura de la horquilla III que incluye los primeros codones de la proteína IcaR o por alguno otro mecanismo desconocido que pueda actuar en *trans* (proteínas de unión a RNA o RNA reguladores pequeños). Cualquiera de estas posibilidades explicarían un posible desacople entre la transcripción y la traducción dando el tiempo suficiente para que el mRNA de *icaR* se transcriba en su totalidad y por lo tanto, la hipótesis de la circularización del mRNA sea válida.

Con respecto a las interacciones intermoleculares entre moléculas de RNA idénticas, a través de secuencias complementarias, se ha visto que son esenciales en determinados mecanismos (Paillart *et al.*, 2004). Por ejemplo estas interacciones se producen durante i) el empaquetamiento de las moléculas de RNA en las cápsides de los retrovirus, ii) en la formación de las partículas de ribonucleoproteínas necesarias para el transporte y localización del mRNA *bcd* (también conocido como *bicoid*) durante el

desarrollo de *Drosophila* (Ferrandon *et al.*, 1997; Wagner *et al.*, 2001), y iii) en la formación del hexámero cíclico pRNA necesario para un empaquetamiento eficiente del genoma del bacteriófago *phi29* de *Bacillus subtilis* (Peixuan Guo *et al.*, 1998; Peixuan Guo, 2005). En estos ejemplos, la interacción intermolecular de los mRNAs proporciona un mecanismo de reclutamiento de las moléculas de mRNA dentro de un complejo estructural específico. El hecho de que la molécula completa del mRNA de *icaR*, y no la 5'-UTR sola, puedan formar dímeros *in vitro* es consistente con la interacción en *trans* entre la 3'-UTR de una molécula de mRNA con la 5'-UTR de otra (Fig. 15). Sin embargo, como los experimentos de expresión ectópica, tanto de la 3'-UTR sola como del mRNA completo de *icaR*, no son capaces de modificar la expresión de la copia cromosómica de *IcaR* cuestiona la relevancia de la interacción intermolecular *in vivo* (datos no mostrados). Además, los resultados obtenidos en los ensayos de *toeprint* y en la digestión por RNase III pueden ser explicados considerando como posibles cualquiera de las interacciones. Es decir, los resultados serían exactamente iguales si las moléculas interactuasen en *trans* o en *cis*. Por lo tanto, es necesario la realización de más estudios que determinen si el mecanismo molecular que controla la expresión de *IcaR* y, en consecuencia, la formación de biofilm en *S. aureus*, ocurre a nivel inter- o intramolecular. Además, otra pregunta que quedaría por resolver es si la interacción entre la 5'- y la 3'-UTR requiere de la participación de otros factores que actuaría en *trans* para alejar la 3'- de la 5'-UTR permitiendo alternar entre una estructura circular cerrada a una abierta que permita el

acceso a la RBS. Aunque en este estudio hemos mostrado que la mutación de dos proteínas de unión a RNA, como son Hfq y RNA helicasa YqfR, no afectan la expresión de IcaR, no podemos excluir que otras proteínas de unión a RNA o pequeños RNA reguladores puedan participar en la interacción entre la 5'- y la 3'-UTR en respuesta a determinadas señales ambientales como la temperatura en este caso.

5.2. Relevancia biológica de la interacción entre la 3'-UTR y la 5'-UTR

Independientemente del mecanismo exacto que gobierna la interacción entre las UTRs, nuestros resultados indican que la habilidad de la 3'-UTR para interactuar con la 5'-UTR tiene profundas consecuencias en la síntesis de la proteína IcaR y en el proceso de formación de biofilm (Fig. 27). Por un lado, el emparejamiento entre estas UTRs genera una región de RNA de cadena doble que es reconocida por la endoribonucleasa RNase III, provocando un aumento en la degradación del mRNA de *icaR*. Por otro lado, el emparejamiento de las UTRs también oculta la RBS y por lo tanto impide la formación del complejo de iniciación de la traducción. En la literatura existen ejemplos similares que muestran que la interacción entre la 3'- y la 5'-UTR modula la expresión proteica en procariotas (Novick *et al.*, 1993; Thisted *et al.*, 1995; Balaban and Novick, 1995; Brenneis and Soppe, 2009). Por ejemplo, en el sistema toxina/antitoxina *hok/sok* del plásmido R1, la 3'-UTR se empareja con la zona de la RBS para inhibir la traducción (Thisted *et al.*, 1995). En consecuencia, la molécula de entera del mRNA de

hok no puede producir la proteína mientras que una versión más corta del mRNA que carece de la región 3' produce la proteína Hok. De igual modo, el grupo de R. Novick proponía un mecanismo de regulación similar donde el extremo 5' del RNAIII interactuaba con el extremo 3' (Novick *et al.*, 1993). Años mas tarde, esta interacción fue demostrada mediante el mapeo enzimático y químico de la estructura secundaria del RNAIII (Benito *et al.*, 2000). Es importante señalar que el RNAIII (514 nt), uno de los RNA reguladores más estudiados en *S. aureus*, es en realidad el mRNA que codifica para δ -hemolisina (26 aa). Esto implica que la principal región reguladora del RNAIII corresponde a la 3'-UTR, la cual tiene una longitud de 352 nt y puede formar una estructura secundaria estable capaz de emparejarse en *trans* con diferentes mRNAs dianas para reprimir su traducción (Felden *et al.*, 2011). Llamativamente, la capacidad del RNAIII como regulador de la traducción de su propia proteína es mucho menos conocida (Balaban and Novick, 1995). Se había demostrado que la delección del extremo 3' del RNAIII suprimía el retraso temporal existente entre la transcripción del RNAIII y su traducción. Aunque el mecanismo no fue explicado en ese momento, se especuló que la interacción entre las UTRs impedía la traducción de la δ -hemolisina y que sería necesario un factor específico adicional que desestructura la interacción para permitir la traducción de dicha proteína (Balaban and Novick, 1995), anticipando de alguna manera los resultados descritos en este estudio.

Una característica común entre los mRNAs de *icaR* y RNAIII es la presencia de motivos UCCCC en sus 3'-UTRs (Huntzinger *et al.*, 2005). Es

interesante notar que este motivo también ha sido encontrado en varios RNA reguladores pequeños de *S. aureus* (Geissmann *et al.*, 2009). En el caso de RNAIII y RsaE, el motivo UCCCC interactúa con mRNAs dianas en *trans* mientras que en *icaR* lo hace con la RBS codificada en la misma molécula de mRNA. Sin embargo, no podemos excluir que el motivo UCCCC de *icaR* interactúe con otras dianas de mRNA. De hecho, las predicciones *in silico* realizadas utilizando la aplicación RNAPredator (Eggenhofer *et al.*, 2011) identifican varios mRNAs candidatos cuyas RBS emparejan con el UCCCC de *icaR* (datos no mostrados). Llamativamente, alguno de estos mRNAs candidatos codifican para proteínas que podrían estar involucradas en el proceso de formación de biofilm como son la N-acetil-glucosamina-transferasa, la proteína TagB necesaria para la biosíntesis de ácidos teicoicos, el transportador dependiente de ATP necesario para la incorporación de espermidina/putrescina o el regulador transcripcional Fur dependiente de hierro. Obviamente, se necesitan estudios adicionales para dilucidar si el motivo UCCCC de *icaR* también puede interactuar en *trans* como RNAIII.

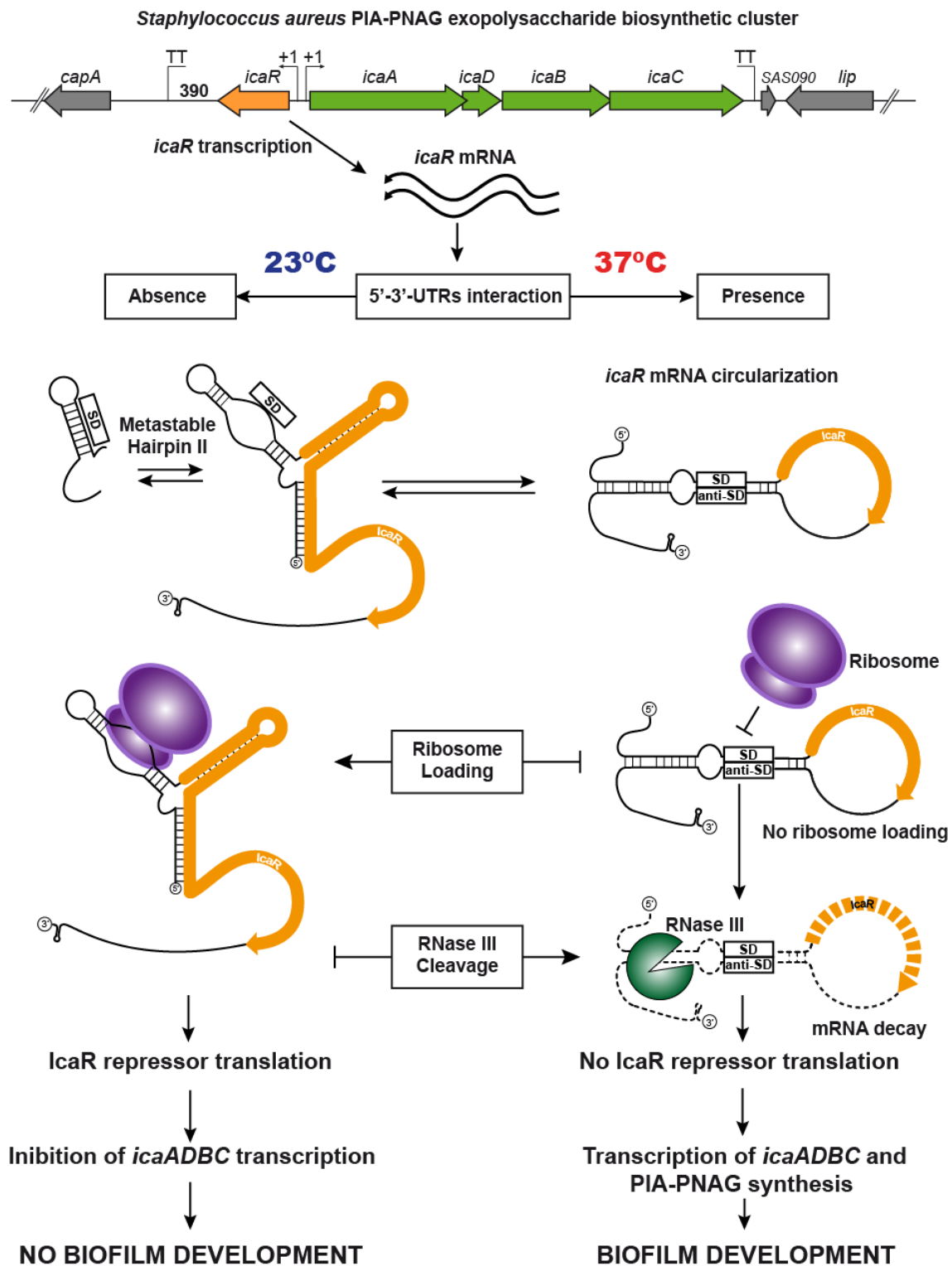


Figure 27. Leyenda en la página siguiente.

Figura 27. Modelo de regulación post-transcripcional de la traducción de IcaR mediante la interacción de la 5'- y la 3'UTR. Una vez que el gen *icaR* comienza a transcribirse, una serie de interacciones RNA-RNA entre varios motivos codificados en la misma molécula de mRNA controla la traducción de la proteína IcaR. Estas interacciones son sensibles a la temperatura y en consecuencia la molécula de mRNA se organiza de manera diferente, según donde la bacteria se encuentre, para controlar la formación del biofilm en el nicho adecuado. A temperatura ambiente (**23°C**), el extremo 5' del mRNA de *icaR* se organiza en una estructura en forma de "Y" que incluye tres horquillas de RNA. Una de las horquillas (hairpin II) es meta-estable alternando entre dos conformaciones, en una de ellas, el dominio *FourU* hibrida con la SD mientras que en la otra la SD se encuentra libre. Esta estructura de tres horquillas no permite el acceso de la 3'-UTR pero sí la traducción de IcaR a 23°C. En cambio a la temperatura del hospedador (**37°C**), la estructura localizada en el extremo 5' sufre un cambio conformacional significativo debido al emparejamiento con la 3'-UTR lo que genera una región de RNA de cadena doble que bloquea la SD a través del dominio UCCCC y promueve el procesamiento por RNase III. El resultado de esta regulación post-transcripcional dependiente de la temperatura es que a temperatura ambiente hay una mayor expresión de IcaR lo que evita la formación de biofilm. En cambio hay una menor expresión de IcaR cuando la bacteria se encuentra en el hospedador, facilitando la síntesis del exopolisacárido PIA-PNAG y así la formación de biofilm.

5.3. La temperatura modula la interacción entre la 3'-UTR y la 5'-UTR del mRNA de *icaR*

La temperatura es uno de los factores ambientales que altera *in vivo* la cantidad final de proteína IcaR y por tanto la producción del biofilm polisacárido en *S. aureus* (Fig. 22). Existen numerosos estudios que demuestran que la temperatura es una señal esencial para la expresión de diversos genes en el contexto ambiental adecuado (Kortmann *et al.*, 2011). Por ejemplo, la expresión de genes de virulencia están controlados a nivel post-transcripcional mediante estructuras secundarias de RNA cuya conformación es sensible a los cambios de temperatura (Johansson *et al.*, 2002; Böhme *et al.*, 2012). En la mayoría de los casos, estas estructuras bloquean el acceso del ribosoma a la región SD cuando el microorganismo patógeno se encuentra a temperatura ambiente. En cambio, cuando dicho patógeno infecta exitosamente a su hospedador, y por lo tanto la

temperatura sube hasta los 37°C, las estructuras secundarias se abren dejando paso al ribosoma, que traducirá los genes necesarios para la supervivencia en el nuevo ambiente. Este tipo de RNA termosensores están formados básicamente por una horquilla de RNA donde uno de sus brazos incluye la región SD, que es rica en Gs, mientras que el otro suele contener una secuencia complementaria que determina la sensibilidad de la horquilla a la temperatura. En algunos casos, en este brazo, hay cuatro Us seguidas que dan el nombre de *FourU* al elemento regulador que hibrida con la SD (Kortmann *et al.*, 2011). Las uniones G-U son más sensibles al incremento de la temperatura, lo que facilita la apertura de la estructura. Estos mecanismo de apertura y cierre de la horquilla de RNA se le conoce con el nombre de *zipper* (o cremallera) (Kortmann *et al.*, 2011).

A pesar de que el mRNA de *icaR* contiene un elemento *FourU* que hibrida con la región SD (Fig. 24C), la expresión de la proteína IcaR dependiente de la temperatura no puede ser explicada por dicho mecanismo *zipper*. Al contrario que la mayoría de los genes regulados por el motivo *FourU*, IcaR se expresa en mayor cantidad a temperatura ambiente (23°C) que a la temperatura del hospedador (37°C). Esto se debe probablemente a que el extremo 5' de la molécula de mRNA de *icaR* contiene más elementos reguladores que una simple horquilla que responde a la temperatura. De hecho, los resultados muestran que la estructura del extremo 5' del mRNA de *icaR* contiene una conformación tipo “Y” con tres horquillas (Fig. 24C). Esta estructura cumple una doble función: por un lado, permite el acceso del ribosoma a baja temperatura y, por el otro, modula la interacción con la

3'-UTR, la cual actúa como elemento regulador negativo de la traducción (Fig. 27). El mapeo químico y enzimático de la estructura del extremo 5' confirma que a bajas temperaturas la conformación de tres horquillas se mantiene cerrada mediante la interacción de una región de la 5'-UTR con otra dentro de la secuencia codificante, tanto en presencia como en ausencia de la región 3'-UTR. Esto explicaría la baja interacción de los fragmentos 5'-UTR y 3'-UTR en los experimentos de retardo (Fig. 23A) y, en consecuencia, la ausencia de procesamiento dependiente de RNase III a 23°C (Fig. 23B). En cambio a 37°C, la estructura sufre un cambio conformacional muy significativo en presencia de la 3'-UTR (Fig. 26B). Las horquillas I y II se abren para interactuar con la región 3'-UTR que provoca la oclusión de la región SD y la digestión por RNase III (Fig. 27). Este drástico cambio conformacional no ocurre en ausencia de la 3'-UTR, lo que indica que es necesario la presencia del extremo 3' para que la estructura se reorganice. En consecuencia, la molécula entera del mRNA de *icaR* actuaría como un RNA *termosensor* que modula la expresión del regulador negativo del biofilm de *S. aureus*. Existe un cierto paralelismo entre este mecanismo y el descrito para el mRNA de *cspA* de *E. coli* (Giuliodori *et al.*, 2010). Del mismo modo que en *icaR*, la estructura del mRNA de *cspA* está organizada de manera que permite la expresión a bajas temperaturas de la proteína CspA, una de las principales proteínas necesarias para la adaptación de la bacteria al estrés por frío en *E. coli* (Giuliodori *et al.*, 2010). En cambio a 37°C, el mRNA sufre un cambio conformacional que genera regiones de RNA de doble cadena que ocultan la región SD para inhibir la

traducción de la proteína a esa temperatura (Giuliodori *et al.*, 2010). De manera general, los mecanismos que implican un cambio conformacional dependiente de la temperatura son conocidos con el nombre de *switch* (cambio) que los diferencia de los *zipper* (Kortmann *et al.*, 2011). En el caso de *cspA*, el cambio conformacional dependiente de la temperatura ocurre en el extremo 5' de la molécula de mRNA, a diferencia de *icaR*, que necesita además de la participación de la 3'-UTR para conseguir la inhibición de la traducción.

El descubrimiento de interacciones RNA-RNA entre dominios codificados en una misma molécula de mRNA y separados por varios cientos de nucleótidos entre ellos, y que a su vez son capaces de medir la temperatura para expresar correctamente una proteína en el entorno adecuado, abre nuevas posibilidades para comprender los procesos de regulación post-transcripcionales en bacterias. Estos nuevos mecanismos de regulación bacteriana mediante 3'-UTRs dependientes de la temperatura se acercan a los descritos en organismos eucariotas. Por ejemplo, al igual que en las bacterias patógenas, la expresión de genes en ciertos parásitos, como *Leishmania*, está controlada de manera post-transcripcional en función de la temperatura. El parásito causante de la leishmaniosis tiene un ciclo de vida alterno entre el mosquito (el vector) y los mamíferos (los hospedadores) y por lo tanto durante el proceso de infección, el parásito tiene que ser capaz de detectar el cambio de temperatura sufrido al pasar del mosquito al hospedador. Esta regulación se da mayoritariamente por elementos reguladores presentes en las 3'-UTRs (Charest *et al.*, 1996;

McNicoll *et al.*, 2005). Uno de los genes específicos que controlan la estancia en el mamífero es la proteína de estrés térmico Hsp83 que se induce a la temperatura de 37°C (Shapira *et al.*, 1988; Zilka *et al.*, 2001; Maya David *et al.*, 2010). La 3'-UTR del mRNA presenta una estructura secundaria rica en pirimidinas que responde ante la temperatura del huésped (Lior David *et al.*, 2006). A 37°C esta estructura se relaja permitiendo la circularización del mRNA necesaria para una correcta traducción de la proteína (Zilka *et al.*, 2001; Maya David *et al.*, 2010). Salvando las diferencias moleculares, este proceso de interacción entre UTRs del mismo mRNA dependiente de la temperatura se asemeja al descrito en esta Tesis Doctoral para el mRNA de *icaR*. Teniendo en cuenta que muchas bacterias patógenas también presentan ciclos de infección alternos entre el ambiente (o un vector) y el hospedador, y del mismo modo, el cambio de temperatura es una de las primeras señales que marcan el cambio de hábitat, sería esperable que este tipo de mecanismos se encuentren extendidos más allá del extremo 5' del mRNA e incluyan estructuras presentes tanto en la ORF como en la 3'-UTR como es el caso de *icaR* en *S. aureus*.

5.4. Los elementos reguladores localizados en las 3'-UTRs estarían presentes en todas las bacterias

¿Por qué las regiones 3'-UTRs han pasado desapercibidas en bacterias? Una posible explicación es que la gran mayoría de los análisis

transcriptómicos realizados hasta ahora en estos organismos proporcionaban muy poca información referente a las características de estas regiones. Esto se debía a que estos análisis estaban focalizados en la identificación de los extremos 5', con el fin de identificar los inicios primarios de transcripción y en consecuencia conocer las regiones promotoras de cada uno (Sharma and Jörg Vogel, 2009; Albrecht *et al.*, 2010; Dornenburg *et al.*, 2010; Mitschke, Georg, *et al.*, 2011; Mitschke, Vioque, *et al.*, 2011; Schmidtke *et al.*, 2012; Wurtzel, Sesto, *et al.*, 2012; Wurtzel, Yoder-Himes, *et al.*, 2012). Además, el hecho de que los genomas bacterianos son más compactos que los eucariotas y la noción de que las 3'-UTRs largas están restringidas a organismos más complejos (Pesole *et al.*, 2002; Mazumder *et al.*, 2003), ha creado la falsa impresión de que los genomas bacterianos no tiene espacio para albergar más que unos pocos RNA reguladores pequeños dentro de sus cortas regiones intergénicas. Sin embargo, este estudio anticipa que las 3'-UTRs bacterianas pueden contener elementos reguladores que contribuyan a la regulación de la expresión génica mediante diferentes mecanismos, incluido probablemente alguno similar a la interacción entre las regiones 3'- y 5' aquí descrita. De hecho recientemente Lopez-Garrido y colaboradores han demostrado que el mRNA de *hilD*, que codifica para el activador transcripcional de la isla de patogenicidad 1 de *Salmonella*, contiene una 3'-UTR larga que es necesaria para la correcta expresión de HilD (Lopez-Garrido *et al.*, 2014). De manera similar que en *icaR*, la 3'-UTR de *hilD* controla el procesamiento

del mRNA. En este caso, la 3'-UTR podría ser diana para la acción del degradosoma (Lopez-Garrido *et al.*, 2014).

Teniendo en cuenta que el análisis del mapa transcriptómico de *S. aureus* confirma que al menos el 35% de los mRNAs de *S. aureus* contiene 3'-UTRs mayores a 100 nt de largo y que las predicciones *in silico* realizadas para otras bacterias también muestran la presencia de 3'-UTRs largas de manera generalizada en el mundo procariota, es esperable que la regulación post-transcripcional mediada por 3'-UTRs sea más relevante para la fisiología bacteriana de lo imaginado hasta ahora. La conservación de secuencia, así como el largo de las 3'-UTRs, son indicadores de que estas regiones pueden albergar elementos adicionales a la señales de terminación de la transcripción. Si analizamos los 1059 terminadores de transcripción Rho-independientes de *S. aureus*, predichos mediante TransTermHP (Kingsford *et al.*, 2007) con una confianza mayor al 90%, encontramos que el largo promedio de estos terminadores es de 33 ± 10 y, en consecuencia, el largo mínimo de una 3'-UTR para albergar la secuencia de terminación debería ser de alrededor de 40-50 nt. Sin embargo, la alta prevalencia de 3'-UTRs largas (>100 nt) es una evidencia clara de que estas regiones cumplen funciones adicionales. La inclusión de elementos reguladores o moduladores de la traducción y/o procesamiento del mRNA en las 3'-UTRs, en lugar de las 5'-UTRs, puede presentar una serie de ventajas debido a que esta región no se ve afectada por el proceso de traducción. Por ello, estas regiones pueden haber sido seleccionadas como diana para los procesos evolutivos bacterianos de manera análoga a lo

descrito en organismos superiores. En relación a esto, es interesante notar que los mRNAs de *icaR* codificados en los genomas de otras especies de estafilococos tales como *S. epidermidis*, *S. simiae*, *S. caprae* y *S. capitis* también contienen 3'-UTRs largas (365, 482, 396 y 380 nt respectivamente). Sin embargo, las secuencias de estas 3'-UTRs no se encuentran conservadas y el motivo *UCCCC* solo está presente en *S. aureus* y *S. simiae*. Esto indica que los mRNAs que codifican para la misma proteína IcaR ha sufrido una evolución diferente en función de la especie bacteriana. Estas diferencias podrían permitir ajustar el proceso de formación del biofilm a un nicho en particular de una manera específica según la especie. La comprensión de cómo distintas 3'-UTRs de genes ortólogos regulan la expresión de las proteínas será una estrategia prometedora para descubrir el potencial regulador de las 3'-UTRs bacterianas y su posible implicación tanto en la especiación bacteriana como en niveles evolutivos más generales.

CONCLUSIONES

6. CONCLUSIONES

1) El análisis del mapa transcriptómico de *S. aureus*, realizado mediante la combinación de técnicas que incluyen microarrays de alta densidad y secuenciación masiva del RNA, ha revelado que un tercio de los mRNAs contienen 3'-UTRs con una longitud superior a los 100 nucleótidos. Esto muestra el potencial de las 3'-UTRs para albergar elementos reguladores en estas regiones.

2) Basados en que la mayoría de los transcritos mapeados en *S. aureus* incluyen un terminador de la transcripción Rho-independiente, y que estos terminadores pueden ser predichos mediante programas informáticos, es posible determinar *in silico* la longitud de las 3'-UTRs en todas las bacterias cuyo genoma se encuentre secuenciado.

3) El análisis bioinformático de 25 genomas procariotas muestra que la longitud de estas 3'-UTRs es muy variable entre las distintas especies bacteriana, e incluso entre las del mismo género, lo que sugiere una posible selección de estas regiones para incluir cambios significativos durante el proceso evolutivo.

4) El análisis de conservación de los nucleótidos en el mRNA de *icaR*, que codifica para el principal regulador transcripcional de la síntesis del exopolisacárido PIA-PNAG necesario para la formación de biofilm en *S. aureus*, revela que la secuencia de la 3'-UTR (390-nt de largo) se encuentra

conservada a niveles similares a los encontrados en la ORF y muy alejados de la variabilidad observada en la región intergénica que no es transcrita. Esta conservación de nucleótidos indica un rol funcional de la 3'-UTR de *icaR*.

5) El análisis genético y bioquímico revela que la 3'-UTR de *icaR* participa en la expresión de la proteína IcaR. La delección de 330-nt de la 3'-UTR (realizada sin afectar el terminador de transcripción) produce un aumento tanto en los niveles del mRNA como en los niveles de la proteína IcaR.

6) La expresión, desde un promotor constitutivo ectópico, de la molécula completa del mRNA de *icaR*, o de su variante que lleva la delección en la 3'-UTR, revela que la proteína IcaR se sigue expresado en mayor cantidad en el mutante 3'-UTR, demostrando que la regulación de la traducción de IcaR mediada por la 3'-UTR ocurre a nivel post-transcripcional.

7) La delección de la 3'-UTR provoca un aumento en la estabilidad del mRNA de *icaR* similar a los observados en un mutante del gen *rnc*, que codifica para la endoribonucleasa de cadena doble RNase III. Esto sugiere la presencia de una región de RNA de cadena doble en el mRNA de *icaR* que es reconocida por dicha enzima.

8) La predicción informática de la estructura secundaria de la molécula completa del mRNA de *icaR* muestra que una región de la 3'-UTR hibrida con casi la totalidad de la 5'-UTR, generando una estructura cuasi-circular del mRNA. Dentro de esta zona de interacción, el motivo UCCCCUG,

localizado 260-nt después del codón stop de la proteína IcaR, hibrida con la secuencia UAGGGGG correspondiente a la región de unión al ribosoma localizada en la 5'-UTR.

9) La interacción entre las regiones 5'- y 3'-UTRs del mRNA de *icaR* promueve el procesamiento del mRNA mediado por RNase III e impide la formación del complejo de iniciación de la traducción. La delección o sustitución del motivo UCCCCUG es suficiente para impedir dicha interacción de dominios, y en consecuencia evitar la digestión mediada por RNase III y permitir el acceso del ribosoma, facilitando la traducción de IcaR.

10) El mayor nivel de la proteína IcaR, provocado tanto por la delección de la 3'-UTR completa como la mutación del motivo UCCCCUG, es suficiente para inhibir de manera completa la formación de biofilm en *S. aureus*, lo que demuestra la relevancia biológica de esta regulación post-transcripcional.

11) Cambios en la temperatura ambiental provocan una expresión diferencial de la proteína IcaR. A 23°C existe una mayor expresión de IcaR que a 37°C, sin que haya variaciones significativas en los niveles del mRNA de *icaR*, lo que indica una regulación post-transcripcional dependiente de la temperatura.

12) El extremo 5' de la molécula de mRNA de *icaR* se organiza en una estructura secundaria en forma de "Y" que incluye tres horquillas de RNA. A

23°C esta estructura está ordenada de manera que permite el acceso del ribosoma y, al mismo tiempo, impide la interacción con la 3'-UTR y en consecuencia la acción de la RNase III. En cambio a 37°C esta estructura sufre un cambio conformacional que permite el emparejamiento con la 3'-UTR, inhibiendo la traducción de IcaR.

13) La región SD (AGGGGG) del mRNA de *icaR* puede interactuar de manera competitiva con al menos dos dominios de RNA codificados en la misma molécula de mRNA: el motivo UCCCC localizado en la 3'-UTR y el dominio *FourU* localizado en la 5'-UTR.

14) La molécula completa del mRNA de *icaR* actuaría como un *RNA termosensor* que cambia su conformación en función de la temperatura para modular el proceso de formación de biofilm según la bacteria se encuentra en el hospedador o fuera de él.

REFERENCIAS

7. REFERENCIAS

Agrawal, N., Dasaradhi, P.V.N., Mohmmmed, A., Malhotra, P., Bhatnagar, R.K., and Mukherjee, S.K. (2003) RNA interference: biology, mechanism, and applications. *Microbiology and Molecular Biology Reviews* **67**: 657–685.

Albrecht, M., Sharma, C.M., Reinhardt, R., Vogel, J., and Rudel, T. (2010) Deep sequencing-based discovery of the *Chlamydia trachomatis* transcriptome. *Nucleic Acids Research* **38**: 868–877.

Altman, S. (2007) A view of RNase P. *Mol BioSyst* **3**: 604–607.

Alvarez, D.E., Lodeiro, M.F., Ludueña, S.J., Pietrasanta, L.I., and Gamarnik, A.V. (2005) Long-range RNA-RNA interactions circularize the dengue virus genome. *Journal of virology* **79**: 6631–6643.

Aly, R., Shinefield, H.R., Litz, C., and Maibach, H.I. (1980) Role of teichoic acid in the binding of *Staphylococcus aureus* to nasal epithelial cells. *CORD Conference Proceedings* **141**: 463–465.

Amarasinghe, A.K., Calin-Jageman, I., Harmouch, A., Sun, W., and Nicholson, A.W. (2001) *Escherichia coli* ribonuclease III: affinity purification of hexahistidine-tagged enzyme and assays for substrate binding and cleavage. *Methods in enzymology* **342**: 143–158.

Anderson, K.L., and Dunman, P.M. (2009) Messenger RNA turnover processes in *Escherichia coli*, *Bacillus subtilis*, and emerging studies in *Staphylococcus aureus*. *International Journal of Microbiology* **2009**: 525491.

Andrade, J.M., Cairrão, F., and Arraiano, C.M. (2006) RNase R affects gene expression in stationary phase: regulation of *ompA*. *Molecular Microbiology* **60**: 219–228.

Arnaud, M., Chastanet, A., and Débarbouillé, M. (2004) New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, Gram-positive bacteria. *Applied and Environmental Microbiology* **70**: 6887–6891.

Arnvig, K.B., Comas, I., Thomson, N.R., Houghton, J., Boshoff, H.I., Croucher, N.J., *et al.* (2011) Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of *Mycobacterium tuberculosis*. *PLoS Pathog* **7**: e1002342.

Arraiano, C.M., Andrade, J.M., Domingues, S., Guinote, I.B., Malecki, M., Matos, R.G., *et al.* (2010) The critical role of RNA processing and

degradation in the control of gene expression. *FEMS Microbiology Reviews* **34(5)**:883-923.

Baba, T., Takeuchi, F., Kuroda, M., Yuzawa, H., Aoki, K.-I., Oguchi, A., *et al.* (2002) Genome and virulence determinants of high virulence community-acquired MRSA. *The Lancet* **359**: 1819–1827.

Babitzke, P., Baker, C.S., and Romeo, T. (2009) Regulation of Translation Initiation by RNA Binding Proteins. *Annu Rev Microbiol* **63**: 27–44.

Balaban, N., and Novick, R.P. (1995) Translation of RNAIII, the *Staphylococcus aureus* agr regulatory RNA molecule, can be activated by a 3'-end deletion. *FEMS Microbiology Letters* **133**: 155–161.

Barkema, H.W., Schukken, Y.H., and Zadoks, R.N. (2006) Invited Review: The role of cow, pathogen, and treatment regimen in the therapeutic success of bovine *Staphylococcus aureus* mastitis. *J Dairy Sci* **89**: 1877–1895.

Baselga, R., Albizu, I., La Cruz, De, M., Del Cacho, E., Barberan, M., and Amorena, B. (1993) Phase variation of slime production in *Staphylococcus aureus*: implications in colonization and virulence. *Infect Immun* **61**: 4857–4862.

Bechhofer, D.H. (2009) Messenger RNA Decay and Maturation in *Bacillus subtilis*. *Prog Mol Biol Transl Sci* **85**: 231–273.

Belasco, J.G. (2010) All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay. *Nat Rev Mol Cell Biol* **11**: 467–478.

Benito, Y., Kolb, F.A., Romby, P., Lina, G., Etienne, J., and Vandenesch, F. (2000) Probing the structure of RNAIII, the *Staphylococcus aureus* agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* **6**: 668–679.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., and Snyder, M. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74.

Böhme, K., Steinmann, R., Kortmann, J., Seekircher, S., Heroven, A.K., Berger, E., Pisano, F., Thiermann, T., Wolf-Watz, H., Narberhaus, F., Dersch, P. (2012) Concerted actions of a thermo-labile regulator and a unique intergenic RNA thermosensor control *Yersinia* virulence. *PLoS Pathog* **8**: e1002518.

Boisset, S., Geissmann, T., Huntzinger, E., Fechter, P., Bendridi, N., Possedko, M., *et al.* (2007) *Staphylococcus aureus* RNAIII coordinately represses the synthesis of virulence factors and the transcription regulator Rot by an antisense mechanism. *Genes & Development* **21**: 1353–1366.

- Boles, B.R., and Horswill, A.R. (2011) Staphylococcal biofilm disassembly. *Trends in Microbiology* **19**: 449–455.
- Boudvillain, M., Figueroa-Bossi, N., and Bossi, L. (2013) Terminator still moving forward: expanding roles for Rho factor. *Current Opinion in Microbiology* **16**: 118–124.
- Bradford, J.R., Hey, Y., Yates, T., Li, Y., Pepper, S.D., and Miller, C.J. (2010) A comparison of massively parallel nucleotide sequencing with oligonucleotide microarrays for global transcription profiling. *BMC Genomics* **11**: 282.
- Breaker, R.R. (2012) Riboswitches and the RNA World. *Cold Spring Harbor Perspectives in Biology* **4**: a003566
- Brenneis, M., and Soppa, J. (2009) Regulation of Translation in Haloarchaea: 5'- and 3'-UTRs Are Essential and Have to Functionally Interact *In Vivo*. *PLoS ONE* **4**: e4484.
- Britton, R.A., Wen, T., Schaefer, L., Pellegrini, O., Uicker, W.C., Mathy, N., *et al.* (2007) Maturation of the 5' end of *Bacillus subtilis* 16S rRNA by the essential ribonuclease YkqC/RNase J1. *Molecular Microbiology* **63**: 127–138.
- Brodersen, P., and Voinnet, O. (2009) Revisiting the principles of microRNA target recognition and mode of action. *Nat Rev Mol Cell Biol* **10**: 141–148.
- Broeke-Smits, ten, N.J.P., Pronk, T.E., Jongerius, I., Bruning, O., Wittink, F.R., Breit, T.M., *et al.* (2010) Operon structure of *Staphylococcus aureus*. *Nucleic Acids Research* **38**: 3263–3274.
- Butcher, S.E., and Pyle, A.M. (2011) The Molecular Interactions That Stabilize RNA Tertiary Structure: RNA Motifs, Patterns, and Networks. *Acc Chem Res* **44**: 1302–1311.
- Cadelari, I., Chao, Y., Romby, P., and Vogel, J. (2013) RNA-mediated regulation in pathogenic bacteria. *Cold Spring Harb Perspect Med* **3**: a010298.
- Calin-Jageman, I., and Nicholson, A.W. (2003) RNA structure-dependent uncoupling of substrate recognition and cleavage by *Escherichia coli* ribonuclease III. *Nucleic Acids Research* **31**: 2381–2392.
- Chambers, H.F., and DeLeo, F.R. (2009) Waves of resistance: *Staphylococcus aureus* in the antibiotic era. *Nat Rev Micro* **7**: 629–641.
- Chambers, H.F., Hartman, B.J., and Tomasz, A. (1985) Increased Amounts of a Novel Penicillin-Binding Protein in a Strain of Methicillin-Resistant *Staphylococcus aureus* Exposed to Nafcillin. *J Clin Invest* **76**: 325–331.

- Chang, Y.-M., Jeng, W.-Y., Ko, T.-P., Yeh, Y.-J., Chen, C.K.-M., and Wang, A.H.-J. (2010) Structural study of TcaR and its complexes with multiple antibiotics from *Staphylococcus epidermidis*. *Proceedings of the National Academy of Sciences* **107**: 8617–8622.
- Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., and Vogel, J. (2012) An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J* **31**: 4005–4019.
- Charest, H., Zhang, W.W., and Matlashewski, G. (1996) The developmental expression of *Leishmania donovani* A2 amastigote-specific genes is post-transcriptionally mediated and involves elements located in the 3'-untranslated region. *J Biol Chem* **271**: 17081–17090.
- Charpentier, E., Anton, A.I., Barry, P., Alfonso, B., Fang, Y., and Novick, R.P. (2004) Novel cassette-based shuttle vector system for Gram-positive bacteria. *Applied and Environmental Microbiology* **70**: 6076–6085.
- Chelladurai, B.S., Li, H., and Nicholson, A.W. (1991) A conserved sequence element in ribonuclease III processing signals is not required for accurate *in vitro* enzymatic cleavage. *Nucleic Acids Research* **19**: 1759–1766.
- Chen, C.Y., Chen, S.T., Juan, H.F., and Huang, H.C. (2012) Lengthening of 3'UTR increases with morphological complexity in animal evolution. *Bioinformatics* **28**: 3178–3181.
- Chen, J.-M., Férec, C., and Cooper, D.N. (2006) A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3' UTR variants. *Hum Genet* **120**: 301–333.
- Cheng, Z.-F., and Deutscher, M.P. (2002) Purification and characterization of the *Escherichia coli* exoribonuclease RNase R. Comparison with RNase II. *J Biol Chem* **277**: 21624–21629.
- Chevalier, C., Boisset, S., Romilly, C., Masquida, B., Fechter, P., Geissmann, T., *et al.* (2010) *Staphylococcus aureus* RNAIII Binds to Two Distant Regions of coa mRNA to Arrest Translation and Promote mRNA Degradation. *PLoS Pathog* **6**: e1000809.
- Condon, C. (2007) Maturation and degradation of RNA in bacteria. *Current Opinion in Microbiology* **10**: 271–278.
- Condon, C., and Putzer, H. (2002) The phylogenetic distribution of bacterial ribonucleases. *Nucleic Acids Research* **30**: 5339–5346.
- Corrigan, R.M., Rigby, D., Handley, P., and Foster, T.J. (2007) The role of *Staphylococcus aureus* surface protein SasG in adherence and biofilm formation. *Microbiology (Reading, Engl)* **153**: 2435–2446.

- Costerton, J.W., Geesey, G.G., and Cheng, K.J. (1978) How bacteria stick. *Sci Am* **238**: 86.
- Costerton, J.W., Stewart, P.S., and Greenberg, E.P. (1999) Bacterial biofilms: a common cause of persistent infections. *Science* **284**: 1318–1322.
- Cramton, S.E., Gerke, C., Schnell, N.F., Nichols, W.W., and Gotz, F. (1999) The intercellular adhesion (*ica*) locus is present in *Staphylococcus aureus* and is required for biofilm formation. *Infect Immun* **67**: 5427–5433.
- Crick, F. (1970) Central dogma of molecular biology. *Nature* **227**: 561–563.
- Croucher, N.J., and Thomson, N.R. (2010) Studying bacterial transcriptomes using RNA-seq. *Current Opinion in Microbiology* **13**: 619–624.
- Cucarella, C., Solano, C., Valle, J., Amorena, B., Lasa, I., and Penades, J.R. (2001) Bap, a *Staphylococcus aureus* surface protein involved in biofilm formation. *Journal of Bacteriology* **183**: 2888–2896.
- David, L., Huber, W., Granovskaia, M., Toedling, J., Palm, C.J., Bofkin, L., et al. (2006) A high-resolution map of transcription in the yeast genome. *PNAS* **103**: 5320–5325.
- David, M., Gabdank, I., Ben-David, M., Zilka, A., Orr, I., Barash, D., and Shapira, M. (2010) Preferential translation of Hsp83 in *Leishmania* requires a thermosensitive polypyrimidine-rich element in the 3' UTR and involves scanning of the 5' UTR. *RNA* **16**: 364–374.
- DeRisi, J., Penland, L., Brown, P.O., and Bittner, M.L. (1996) Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nature*. **14(4)**:457-60.
- Desnoyers, G., Morissette, A., Prévost, K., and Massé, E. (2009) Small RNA-induced differential degradation of the polycistronic mRNA *iscRSUA*. *EMBO J* **28**: 1551–1561.
- Deutscher, M.P. (2005) Degradation of RNA in bacteria: comparison of mRNA and stable RNA. *Nucleic Acids Research* **34**: 659–666.
- Dinges, M.M., Orwin, P.M., and Schlievert, P.M. (2000) Exotoxins of *Staphylococcus aureus*. *Clinical Microbiology Reviews* **13**: 16–34
- Dornenburg, J.E., Devita, A.M., Palumbo, M.J., and Wade, J.T. (2010) Widespread antisense transcription in *Escherichia coli*. *mBio* **1**. e00024-10
- Draper, D.E. (2013) Folding of RNA tertiary structure: Linkages between backbone phosphates, ions, and water. *Biopolymers* **99**: 1105–1113.

- Eggenhofer, F., Tafer, H., Stadler, P.F., and Hofacker, I.L. (2011) RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Research* **39**: W149–54.
- Evguenieva-Hackenberg, E., and Klug, G. (2000) RNase III processing of intervening sequences found in helix 9 of 23S rRNA in the alpha subclass of Proteobacteria. *Journal of Bacteriology* **182**: 4719–4729.
- Fechter, P., Chevalier, C., Yusupova, G., Yusupov, M., Romby, P., and Marzi, S. (2009) Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods in molecular biology (Clifton, NJ)* **540**: 247–263.
- Felden, B., Vandenesch, F., Bouloc, P., and Romby, P. (2011) The *Staphylococcus aureus* RNome and Its Commitment to Virulence. *PLoS Pathog* **7**(3): e1002006.
- Ferrandon, D., Koch, I., Westhof, E., and Nüsslein-Volhard, C. (1997) RNA-RNA interaction is required for the formation of specific *bicoid* mRNA 3' UTR-STAUEN ribonucleoprotein particles. *EMBO J* **16**: 1751–1758.
- Fitzpatrick, F., Humphreys, H., and O'Gara, J.P. (2005) The genetics of staphylococcal biofilm formation-will a greater understanding of pathogenesis lead to better management of device-related infection? *Clin Microbiol Infect* **11**: 967–973.
- Flemming, H.-C., and Wingender, J. (2010) The biofilm matrix. *Nat Rev Micro* **8**: 623–633.
- Foster, T.J. (2005) Immune evasion by staphylococci. *Nat Rev Micro* **3**: 948–958.
- Foster, T.J., and Hook, M. (1998) Surface protein adhesins of *Staphylococcus aureus*. *Trends in Microbiology* **6**: 484–488.
- Frazão, C.C., McVey, C.E.C., Amblar, M.M., Barbas, A.A., Vonnrhein, C.C., Arraiano, C.M.C., and Carrondo, M.A.M. (2006) Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature* **443**: 110–114.
- Fröhlich, K.S., and Vogel, J. (2009) Activation of gene expression by small RNA. *Current Opinion in Microbiology* **12**: 674–682.
- Geissmann, T., Chevalier, C., Cros, M.-J., Boisset, S., Fechter, P., Noirot, C., et al. (2009) A search for small noncoding RNAs in *Staphylococcus aureus* reveals a conserved sequence motif for regulation. *Nucleic Acids Research* **37**: 7239–7257.

- Georg, J., and Hess, W.R. (2011) *cis*-antisense RNA, another level of gene regulation in bacteria. *Microbiol Mol Biol Rev* **75**: 286–300.
- Gerke, C. (1998) Characterization of the *N*-Acetylglucosaminyltransferase Activity Involved in the Biosynthesis of the *Staphylococcus epidermidis* Polysaccharide Intercellular Adhesin. *Journal of Biological Chemistry* **273**: 18586–18593.
- Ghigo, J.M. (2001) Natural conjugative plasmids induce bacterial biofilm development. *Nature* **412**: 442–445.
- Ghosh, S., and Deutscher, M.P. (1999) Oligoribonuclease is an essential component of the mRNA decay pathway. *PNAS* **96**: 4372–4377.
- Giuliodori, A.M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., *et al.* (2010) The *cspA* mRNA Is a Thermosensor that Modulates Translation of the Cold-Shock Protein CspA. *Molecular Cell* **37**: 21–33.
- Gordon, R.J., and Lowy, F.D. (2008) Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. *Clinical Infectious Diseases* **46 Suppl 5**: S350–9.
- Gottesman, S., and Storz, G. (2011) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harbor Perspectives in Biology* **3(12)**: a003798
- Götz, F. (2002) *Staphylococcus* and biofilms. *Molecular Microbiology* **43**: 1367–1378.
- Guo, L., Allen, E.M., and Miller, W.A. (2001) Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. *Molecular Cell* **7**: 1103–1109.
- Guo, P. (2005) RNA nanotechnology: engineering, assembly and applications in detection, gene delivery and therapy. *J Nanosci Nanotechnol* **5**: 1964–1982.
- Guo, P., Zhang, C., Chen, C., Garver, K., and Trottier, M. (1998) Inter-RNA Interaction of Phage ϕ 29 pRNA to Form a Hexameric Complex for Viral DNA Transportation. *Molecular Cell* **2**: 149–155.
- Halees, A.S., Hitti, E., Al-Saif, M., Mahmoud, L., Vlasova-St Louis, I.A., Beisang, D.J., *et al.* (2011) Global assessment of GU-rich regulatory content and function in the human transcriptome. *RNA biology* **8**: 681–691.
- Hall, A.E., Domanski, P.J., Patel, P.R., Vernachio, J.H., Syribey, P.J., Gorovits, E.L., *et al.* (2003) Characterization of a Protective Monoclonal Antibody Recognizing *Staphylococcus aureus* MSCRAMM Protein Clumping Factor A. *Infect Immun* **71**: 6864–6870.

- Hanke, M.L., and Kielian, T. (2012) Deciphering mechanisms of staphylococcal biofilm evasion of host immunity. *Front Cell Infect Microbiol* **2**.
- Hanssen, A.-M., and Ericson Sollid, J.U. (2006) SCCmec in staphylococci: genes on the move. *FEMS Immunology & Medical Microbiology* **46**: 8–20.
- He, J., Jing, S., Zhang, M., Liu, X., Chen, Y., Qie, B., *et al.* (2012) The Measurement of rho-independent Transcription Terminator Efficiency. *1*:1721.1/73911
- He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., and Kinzler, K.W. (2008) The Antisense Transcriptomes of Human Cells. *Science* **322**: 1855–1857.
- Heilmann, C., Schweitzer, O., Gerke, C., Vanittanakom, N., Mack, D., and Gotz, F. (1996) Molecular basis of intercellular adhesion in the biofilm-forming *Staphylococcus epidermidis*. *Molecular Microbiology* **20**: 1083–1091.
- Henz, S.R.S., Cumbie, J.S.J., Kasschau, K.D.K., Lohmann, J.U.J., Carrington, J.C.J., Weigel, D.D., and Schmid, M.M. (2007) Distinct expression patterns of natural antisense transcripts in Arabidopsis. *Plant physiology* **144**: 1247–1255.
- Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., *et al.* (2005) *Staphylococcus aureus* RNAlII and the endoribonuclease III coordinately regulate *spa* gene expression. *EMBO J* **24**: 824–835.
- Høvik, H., Yu, W.-H., Olsen, I., and Chen, T. (2012) Comprehensive transcriptome analysis of the periodontopathogenic bacterium *Porphyromonas gingivalis* W83. *Journal of Bacteriology* **194**: 100–114.
- Itaya, M., Omori, A., Kanaya, S., Crouch, R.J., Tanaka, T., and Kondo, K. (1999) Isolation of RNase H genes that are essential for growth of *Bacillus subtilis* 168. *Journal of Bacteriology* **181**: 2118–2123.
- Izano, E.A., Amarante, M.A., Kher, W.B., and Kaplan, J.B. (2008) Differential roles of poly-N-acetylglucosamine surface polysaccharide and extracellular DNA in *Staphylococcus aureus* and *Staphylococcus epidermidis* biofilms. *Applied and Environmental Microbiology* **74**: 470–476.
- Jefferson, K.K., Cramton, S.E., Götz, F., and Pier, G.B. (2003) Identification of a 5-nucleotide sequence that controls expression of the *ica* locus in *Staphylococcus aureus* and characterization of the DNA-binding properties of IcaR. *Molecular Microbiology* **48**: 889–899.
- Jeng, W.-Y., Ko, T.-P., Liu, C.-I., Guo, R.-T., Liu, C.-L., Shr, H.-L., and

- Wang, A.H.-J. (2008) Crystal structure of IcaR, a repressor of the TetR family implicated in biofilm formation in *Staphylococcus epidermidis*. *Nucleic Acids Research* **36**: 1567–1577.
- Jevons, M.P., Rolinson, G.N., and Knox, R. (1961) Celbenin-Resistant *Staphylococci*. *Br Med J* **1**(5219): 125–126.
- Joanne, P., Falord, M., Chesneau, O., Lacombe, C., Castano, S., Desbat, B., *et al.* (2009) Comparative Study of Two Plasticins: Specificity, Interfacial Behavior, and Bactericidal Activity. *Biochemistry* **48**: 9372–9383.
- Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M., and Cossart, P. (2002) An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. *Cell* **110**: 551–561.
- Jones, R.N. (2008) Key considerations in the treatment of complicated staphylococcal infections. *Clin Microbiol Infect* **14**: 3-9
- Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Dutttagupta, R., Willingham, A.T., *et al.* (2007) RNA Maps Reveal New RNA Classes and a Possible Function for Pervasive Transcription. *Science* **316**: 1484–1488.
- Kawano, M. (2005) Detection of 5' and 3'UTR derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Research* **33**: 1040–1050.
- Khabar, K.S.A. (2010) Post-transcriptional control during chronic inflammation and cancer: a focus on AU-rich elements. *Cell Mol Life Sci* **67**: 2937–2955.
- Kingsford, C.L., Ayanbule, K., and Salzberg, S.L. (2007) Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. *Genome Biol* **8**: R22.
- Klein, D.J., and Ferré-D'Amaré, A.R. (2006) Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science* **313**: 1752–1756.
- Kortmann, J., and Narberhaus, F. (2012) Bacterial RNA thermometers: molecular zippers and switches. *Nature Publishing Group* **10**: 255–265.
- Kortmann, J., Sczodrok, S., Rinnenthal, J., Schwalbe, H., and Narberhaus, F. (2011) Translation on demand by a simple RNA-based thermosensor. *Nucleic Acids Research* **39**: 2855–2868.
- Kuroda, M., Ito, R., Tanaka, Y., Yao, M., Matoba, K., Saito, S., *et al.* (2007) *Staphylococcus aureus* surface protein SasG contributes to intercellular autoaggregation of *Staphylococcus aureus*. *Biochemical and Biophysical Research Communications* **377**: 1102–1106.

- Labandeira-Rey, M., Couzon, F., Boisset, S., Brown, E.L., Bes, M., Benito, Y., *et al.* (2007) *Staphylococcus aureus* Pantón-Valentine Leukocidin Causes Necrotizing Pneumonia. *Science* **315**: 1130–1133.
- Laederach, A., Das, R., Vicens, Q., Pearlman, S.M., Brenowitz, M., Herschlag, D., and Altman, R.B. (2007) Semiautomated and rapid quantification of nucleic acid footprinting and structure mapping experiments. *Nat Protoc* **3**: 1395–1401.
- Lai, D., Proctor, J.R., and Meyer, I.M. (2013) On the importance of cotranscriptional RNA structure formation. *RNA* **19**: 1461–1473.
- Lasa, I., Toledo-Arana, A., and Gingeras, T.R. (2012) An effort to make sense of antisense transcription in bacteria. *RNA biology* **9**: 1039–1044.
- Lasa, I., Toledo-Arana, A., Dobin, A., Villanueva, M., Ruiz de Los Mozos, I., Vergara-Irigaray, M., *et al.* (2011) Genome-wide antisense transcription drives mRNA processing in bacteria. *PNAS* **108**: 20172–20177.
- Lashkari, D.A., DeRisi, J.L., McCusker, J.H., Namath, A.F., Gentile, C., Hwang, S.Y., *et al.* (1997) Yeast microarrays for genome wide parallel genetic and gene expression analysis. *PNAS* **94**: 13057–13062.
- Lee, E.-J., and Groisman, E.A. (2010) An antisense RNA that governs the expression kinetics of a multifunctional virulence gene. *Molecular Microbiology* **76**: 1020–1033.
- Leonard, F.C., and Markey, B.K. (2008) Meticillin-resistant *Staphylococcus aureus* in animals: A review. *The Veterinary Journal* **175**: 27–36.
- Lescoute, A., and Westhof, E. (2005) Topology of three-way junctions in folded RNAs. *RNA* **12**: 83–93.
- Li, Y., and Altman, S. (2003) A specific endoribonuclease, RNase P, affects gene expression of polycistronic operon mRNAs. *PNAS* **100**: 13213–13218.
- Li, Z., Pandit, S., and Deutscher, M.P. (1999) RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO J.* **18**(10): 2878–2885
- Li, Z.W., and Deutscher, M.P. (2002) RNase E plays an essential role in the maturation of *Escherichia coli* tRNA precursors. *RNA* **8**: 97–109.
- Lin, P.-H., and Lin-Chao, S. (2005) RhlB helicase rather than enolase is the beta-subunit of the *Escherichia coli* polynucleotide phosphorylase (PNPase)-exoribonucleolytic complex. *PNAS* **102**: 16590–16595.
- Lin-Chao, S., Wei, C.L., and Lin, Y.T. (1999) RNase E is required for the maturation of ssrA RNA and normal ssrA RNA peptide-tagging activity.

PNAS **96**: 12406–12411.

Lioliou, E., Romilly, C., and Geissmann, T. (2012) RNA-mediated regulation of virulence gene expression: Another layer of complexity. In *Bacterial Regulatory Networks*. Filloux, A.A.M. (ed.). Bacterial Regulatory Networks. 143–165

Lioliou, E., Sharma, C.M., Caldelari, I., Helfer, A.C., Fechter, P., Vandenesch, F., *et al.* (2012) Global regulatory functions of the *Staphylococcus aureus* endoribonuclease III in gene expression. *PLoS Genet* **8**: e1002782.

Liou, G.-G., Chang, H.-Y., Lin, C.-S., and Lin-Chao, S. (2002) DEAD box RhlB RNA helicase physically associates with exoribonuclease PNPase to degrade double-stranded RNA independent of the degradosome-assembling region of RNase E. *J Biol Chem* **277**: 41157–41162.

Liu, F., and Altman, S. (2009) *Ribonuclease P*. Protein Reviews. Series Editor M. Zouhair Atassi. Springer.

Liu, J.M., and Camilli, A. (2010) A broadening world of bacterial small RNAs. *Current Opinion in Microbiology* **13**: 18–23.

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., *et al.* (2009) A trans-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in *Listeria monocytogenes*. *Cell* **139**: 770–779.

López-Garrido, J., Puerta-Fernández, E., and Casadesús, J. (2014) A eukaryotic-like 3' untranslated region in *Salmonella enterica* *hlyD* mRNA. *Nucleic Acids Research* **42**: 5894–5906.

Lowy, F.D. (1998) *Staphylococcus aureus* Infections. *N Engl J Med* **339**: 520–532.

Lundberg, U., and Altman, S. (1995) Processing of the precursor to the catalytic RNA subunit of RNase P from *Escherichia coli*. *RNA* **1**: 327–334.

Lusvarghi, S., Sztuba-Solinska, J., Purzycka, K.J., Rausch, J.W., and Le Grice, S.F.J. (2012) RNA secondary structure prediction using high-throughput SHAPE. *J Vis Exp* e50243–e50243.

Lybecker, M., Zimmermann, B., Bilusic, I., Tukhtubaeva, N., and Schroeder, R. (2014) The double-stranded transcriptome of *Escherichia coli*. *PNAS*. **111**: 3134–3139

Mack, D., Fischer, W., Krokotsch, A., Leopold, K., Hartmann, R., Egge, H., and Laufs, R. (1996) The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: Purification and structural analysis. *Journal of*

Bacteriology **178**: 175–183.

Maher, C.A., Kumar-Sinha, C., Cao, X., Kalyana-Sundaram, S., Han, B., Jing, X., *et al.* (2009) Transcriptome sequencing to detect gene fusions in cancer. *Nature* **458**: 97–101.

Maira-Litrán, T., Kropec, A., Goldmann, D.A., and Pier, G.B. (2005) Comparative opsonic and protective activities of *Staphylococcus aureus* conjugate vaccines containing native or deacetylated Staphylococcal Poly-N-acetyl-beta-(1-6)-glucosamine. *Infect Immun* **73**: 6752–6762.

Masse, E. (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *PNAS* **99**: 4620–4625.

Massé, E., Salvail, H., Desnoyers, G., and Arguin, M. (2007) Small RNAs controlling iron metabolism. *Current Opinion in Microbiology* **10**: 140–145.

Mathy, N., Bénard, L., Pellegrini, O., Daou, R., and Wen, T. (2007) 5'-to-3' Exoribonuclease Activity in Bacteria: Role of RNase J1 in rRNA Maturation and 5' Stability of mRNA. *Cell*. **129**: 681-92.

Matoulkova, E., Michalova, E., Vojtesek, B., and Hrstka, R. (2012) The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA biology* **9**: 563–576.

Mattick, J.S. (2005) The Functional Genomics of Noncoding RNA. *Science* **309**: 1527–1528.

Mayr, C., and Bartel, D.P. (2009) Widespread Shortening of 3' UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell* **138**: 673–684.

Mazumder, B., Seshadri, V., and Fox, P.L. (2003) Translational control by the 3'-UTR: the ends specify the means. *Trends in Biochemical Sciences* **28**: 91–98.

McKenney, D., Hübner, J., Muller, E., Wang, Y., Goldmann, D.A., and Pier, G.B. (1998) The *ica* locus of *Staphylococcus epidermidis* encodes production of the capsular polysaccharide/adhesin. *Infect Immun* **66**: 4711–4720.

McNicoll, F., Muller, M., Cloutier, S., Boilard, N., Rochette, A., Dube, M., and Papadopoulos, B. (2005) Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in *Leishmania*. *J Biol Chem* **280**: 35238–35246.

Merino, E.J., Wilkinson, K.A., Coughlan, J.L., and Weeks, K.M. (2005) RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc* **127**: 4223–4231.

- Merino, N., Toledo-Arana, A., Vergara-Irigaray, M., Valle, J., Solano, C., Calvo, E., *et al.* (2009) Protein A-Mediated Multicellular Behavior in *Staphylococcus aureus*. *Journal of Bacteriology* **191**: 832–843.
- Mignone, F., Gissi, C., Liuni, S., and Pesole, G. (2002) Untranslated regions of mRNAs. *Genome Biol* **3**: 1–10.
- Misra, T.K., and Apirion, D. (1979) RNase E, an RNA processing enzyme from *Escherichia coli*. *J Biol Chem* **254**: 11154–11159.
- Mitra, A., Kesarwani, A.K., Pal, D., and Nagaraja, V. (2010) WebGeSTer DB-a transcription terminator database. *Nucleic Acids Research* **39**: D129–D135.
- Mitschke, J., Georg, J., Scholz, I., Sharma, C.M., Dienst, D., Bantscheff, J., *et al.* (2011) An experimentally anchored map of transcriptional start sites in the model cyanobacterium *Synechocystis* sp. PCC6803. **108**: 2124–2129.
- Mitschke, J., Vioque, A., Haas, F., Hess, W.R., and Muro-Pastor, A.M. (2011) Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. *PNAS* **108**: 20130–20135.
- Mockler, T.C., Chan, S., Sundaresan, A., Chen, H., Jacobsen, S.E., and Ecker, J.R. (2005) Applications of DNA tiling arrays for whole-genome analysis. *Genomics* **85**: 1–15.
- Montanaro, L., Speziale, P., Campoccia, D., Ravaioli, S., Cangini, I., Pietrocola, G., *et al.* (2011) Scenery of *Staphylococcus* implant infections in orthopedics. *Future Microbiology* **6**: 1329–1349.
- Morozova, O., Hirst, M., and Marra, M.A. (2009) Applications of New Sequencing Technologies for Transcriptome Analysis. *Annu Rev Genom Human Genet* **10**: 135–151.
- Mulcahy, M.E., Geoghegan, J.A., Monk, I.R., O'Keeffe, K.M., Walsh, E.J., Foster, T.J., and McLoughlin, R.M. (2012) Nasal Colonisation by *Staphylococcus aureus* Depends upon Clumping Factor B Binding to the Squamous Epithelial Cell Envelope Protein Loricrin. *PLoS Pathog* **8**: e1003092
- Møller-Jensen, J., Franch, T., and Gerdes, K. (2001) Temporal translational control by a metastable RNA structure. *J Biol Chem* **276**: 35707–35713.
- Narberhaus, F. (2010) Translational control of bacterial heat shock and virulence genes by temperature-sensing mRNAs. *RNA biology* **7**: 84–89.
- Nashimoto, M. (1997) Distribution of both lengths and 5' terminal nucleotides of mammalian pre-tRNA 3' trailers reflects properties of 3'

processing endoribonuclease. *Nucleic Acids Research* **25**: 1148-1154.

Nickerson, S.C. (2009) Control of heifer mastitis: Antimicrobial treatment-An overview. *Veterinary Microbiology* **134**: 128–135.

Noller, H.F. (2005) RNA structure: reading the ribosome. *Science* **309**: 1508–1514.

Novick, R.P. (2003) Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Molecular Microbiology* **48**: 1429–1449.

Novick, R.P., Ross, H.F., Projan, S.J., Kornblum, J., Kreiswirth, B., and Moghazeh, S. (1993) Synthesis of staphylococcal virulence factors is controlled by a regulatory RNA molecule. *EMBO J* **12**: 3967–3975.

O'Gara, J.P. (2007) *ica* and beyond: biofilm mechanisms and regulation in *Staphylococcus epidermidis* and *Staphylococcus aureus*. *FEMS Microbiology Letters* **270**: 179–188.

O'Neill, E., Pozzi, C., Houston, P., Humphreys, H., Robinson, D.A., Loughman, A., et al. (2008) A novel *Staphylococcus aureus* biofilm phenotype mediated by the fibronectin-binding proteins, FnBPA and FnBPB. *Journal of Bacteriology* **190**: 3835–3850.

Ogston, A. (1881) Report upon Micro-Organisms in Surgical Diseases. *Br Med J* **1**: 369–375.

Okada, Y., Wachi, M., Hirata, A., Suzuki, K., Nagai, K., and Matsushashi, M. (1994) Cytoplasmic Axial Filaments in *Escherichia Coli* Cells Possible Function in the Mechanism of Chromosome Segregation and Cell-Division. *Journal of Bacteriology* **176**: 917–922.

Okoniewski, M.J., and Miller, C.J. (2006) Hybridization interactions between probesets in short oligo microarrays lead to spurious correlations. *BMC Bioinformatics* **7**: 276.

Otto, M. (2009) *Staphylococcus epidermidis* the “accidental” pathogen. *Nat Rev Micro* **7**: 555–567.

Paillart, J.C., Shehu-Xhilaga, M., Marquet, R., and Mak, J. (2004) Dimerization of retroviral RNA genomes: An inseparable pair. *Nat Rev Micro* **2**: 461–472.

Passalacqua, K.D., Varadarajan, A., Ondov, B.D., Okou, D.T., Zwick, M.E., and Bergman, N.H. (2009) Structure and Complexity of a Bacterial Transcriptome. *Journal of Bacteriology* **191**: 3203–3211.

Pellegrini, O., Nezzar, J., Marchfelder, A., Putzer, H., and Condon, C. (2003) Endonucleolytic processing of CCA-less tRNA precursors by RNase

Z in *Bacillus subtilis*. *EMBO J* **22**: 4534–4543.

Pena, M.D.L., Dufour, D., and Gallego, J. (2009) Three-way RNA junctions with remote tertiary contacts: A recurrent and highly versatile fold. *RNA* **15**: 1949–1964.

Perocchi, F., Xu, Z., Clauder-Munster, S., and Steinmetz, L.M. (2007) Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Research* **35**: e128.

Pesole, G., Liuni, S., Grillo, G., Licciulli, F., Mignone, F., Gissi, C., and Saccone, C. (2002) UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* **30**: 335–340.

Peters, J.M., Vangeloff, A.D., and Landick, R. (2011) Bacterial Transcription Terminators: The RNA 3'-End Chronicles. *Journal of Molecular Biology* **412**: 793–813.

Pinchuk, I.V., Beswick, E.J., and Reyes, V.E. (2010) Staphylococcal enterotoxins. *Toxins (Basel)* **2**: 2177–2197.

Rachid, S., Ohlsen, K., Wallner, U., Hacker, J., Hecker, M., and Ziebuhr, W. (2000) Alternative Transcription Factor sigma B Is Involved in Regulation of Biofilm Expression in a *Staphylococcus aureus* Mucosal Isolate. *Journal of Bacteriology* **182**: 6824–6826.

Ramos, J.L., Martínez-Bueno, M., Molina-Henares, A.J., Terán, W., Watanabe, K., Zhang, X., et al. (2005) The TetR family of transcriptional repressors. *Microbiology and Molecular Biology Reviews* **69**: 326–356.

Rasmussen, S., Nielsen, H.B., and Jarmer, H. (2009) The transcriptionally active regions in the genome of *Bacillus subtilis*. *Molecular Microbiology* **73**: 1043–1057.

Régnier, P., and Hajnsdorf, E. (2013) The interplay of Hfq, poly(A) polymerase I and exoribonucleases at the 3 ends of RNAs resulting from Rho-independent termination: A tentative model. *RNA biology* **10**: 602–609.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., and Pachter, L. (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* **12**: R22.

Romby, P., and Charpentier, E. (2009) An overview of RNAs with regulatory functions in gram-positive bacteria. *Cell Mol Life Sci* **67**: 217–237.

Royce, T.E., Rozowsky, J.S., and Gerstein, M.B. (2007) Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification. *Nucleic Acids Research* **35**: e99.

- Salysers, A.A., and Whitt, D.D. (2011) Bacterial Pathogenesis: A Molecular Approach Amer Society for Microbiology Series. ASM Press, 2011 - 526
- Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., and Burge, C.B. (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* **320**: 1643–1647.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995) Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray. *Science* **270**: 467–470.
- Schmidtke, C., Findeiss, S., Sharma, C.M., Kuhfuss, J., Hoffmann, S., Vogel, J., *et al.* (2012) Genome-wide transcriptome analysis of the plant pathogen *Xanthomonas* identifies sRNAs with putative virulence functions. *Nucleic Acids Research* **40**: 2020–2031.
- Segura, V., Toledo-Arana, A., Uzqueda, M., Lasa, I., and Muñoz-Barrutia, A. (2012) Wavelet-based detection of transcriptional activity on a novel *Staphylococcus aureus* tiling microarray. *BMC Bioinformatics* **13**: 222.
- Serganov, A., and Nudler, E. (2013) A decade of riboswitches. *Cell* **152**: 17–24.
- Serganov, A., Rak, A., Garber, M., Reinbolt, J., Ehresmann, B., Ehresmann, C., *et al.* (1997) Ribosomal protein S15 from *Thermus thermophilus*. Cloning, sequencing, overexpression of the gene and RNA-binding properties of the protein. *European journal of biochemistry / FEBS* **246**: 291–300.
- Sesto, N., Wurtzel, O., Archambaud, C., Sorek, R., and Cossart, P. (2013) The excludon: a new concept in bacterial antisense RNA-mediated gene regulation. *Nat Rev Micro* **11**: 75–82.
- Shanks, R.M.Q., Meehl, M.A., Brothers, K.M., Martinez, R.M., Donegan, N.P., Graber, M.L., *et al.* (2008) Genetic Evidence for an Alternative Citrate-Dependent Biofilm Formation Pathway in *Staphylococcus aureus* That Is Dependent on Fibronectin Binding Proteins and the GraRS Two-Component Regulatory System. *Infect Immun* **76**: 2469–2477.
- Shapira, M., McEwen, J.G., and Jaffe, C.L. (1988) Temperature effects on molecular processes which lead to stage differentiation in *Leishmania*. *EMBO J* **7**: 2895–2901.
- Sharma, C., and Heidrich, N. (2012) Small RNAs and virulence in bacterial pathogens. *RNA biology* **9**: 361–363.
- Sharma, C.M., and Vogel, J. (2009) Experimental approaches for the discovery and characterization of regulatory small RNA. *Current Opinion in Microbiology* **12**: 536–546.

- Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., *et al.* (2008) Deep Sequencing Analysis of Small Noncoding RNA and mRNA Targets of the Global Post-Transcriptional Regulator, Hfq. *PLoS Genet* **4**.
- Spickler, C., and Mackie, G.A. (2000) Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *Journal of Bacteriology* **182**: 2422–2427.
- Spink, W.W. (1954) Staphylococcal infections and the problem of antibiotic-resistant Staphylococci. *Arch Intern Med* **94**: 167.
- Stoodley, P., Sauer, K., Davies, D.G., and Costerton, J.W. (2002) Biofilms as complex differentiated communities. *Annu Rev Microbiol* **56**: 187–209.
- Tadokoro, T., and Kanaya, S. (2009) Ribonuclease H: molecular diversities, substrate binding domains, and catalytic mechanism of the prokaryotic enzymes. *FEBS Journal* **276**: 1482–1493.
- Thisted, T., Sorensen, N.S., and Gerdes, K. (1995) Mechanism of post-segregational killing: Secondary structure analysis of the entire Hok mRNA from plasmid R1 suggests a fold-back structure that prevents translation and antisense RNA binding. *Journal of Molecular Biology* **247**: 859–873.
- Thomason, M.K., and Storz, G. (2010) Bacterial Antisense RNAs: How Many Are There, and What Are They Doing? *. *Annu Rev Genet* **44**: 167–188.
- Thompson, J.F., and Milos, P.M. (2011) The properties and applications of single-molecule DNA sequencing. *Genome Biol* **12**: 217.
- Tijerina, P., Mohr, S., and Russell, R. (2007) DMS footprinting of structured RNAs and RNA-protein complexes. *Nat Protoc* **2**: 2608–2623.
- Toledo-Arana, A., and Solano, C. (2010) Deciphering the physiological blueprint of a bacterial cell. *Bioessays* **32**: 461–467.
- Toledo-Arana, A., Dussurget, O., Nikitas, G., Sesto, N., Guet-Revillet, H., Balestrino, D., *et al.* (2009) The *Listeria* transcriptional landscape from saprophytism to virulence. *Nature* **459**: 950–956.
- Toledo-Arana, A., Merino, N., Vergara-Irigaray, M., Débarbouillé, M., Penadés, J.R., and Lasa, I. (2005) *Staphylococcus aureus* develops an alternative, *ica*-independent biofilm in the absence of the *arlRS* two-component system. *Journal of Bacteriology* **187**: 5318–5329.
- Toledo-Arana, A., Repoila, F., and Cossart, P. (2007) Small noncoding RNAs controlling pathogenesis. *Current Opinion in Microbiology* **10**: 182–188.

- Tormo, M.A., Marti, M., Valle, J., Manna, A.C., Cheung, A.L., Lasa, I., and Penades, J.R. (2005) SarA is an essential positive regulator of *Staphylococcus epidermidis* biofilm development. *Journal of Bacteriology* **187**: 2348–2356.
- Udekwi, K.I., Darfeuille, F., Vogel, J., Reimegård, J., Holmqvist, E., and Wagner, E.G.H. (2005) Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes & Development* **19**: 2355–2366.
- Úbeda, C., Barry, P., Penadés, J.R., and Novick, R.P. (2007) A pathogenicity island replicon in *Staphylococcus aureus* replicates as an unstable plasmid. *PNAS* **104**: 14182–14188.
- Valle, J., Solano, C., García, B., Toledo-Arana, A., and Lasa, I. (2013) Biofilm switch and immune response determinants at early stages of infection. *Trends in Microbiology* **21**: 364–371.
- Valle, J., Toledo-Arana, A., Berasain, C., Ghigo, J.-M., Amorena, B., Penadés, J.R., and Lasa, I. (2003) SarA and not sigma B is essential for biofilm development by *Staphylococcus aureus*. *Molecular Microbiology* **48**: 1075–1087.
- van Vliet, A.H., and Wren, B.W. (2009) New levels of sophistication in the transcriptional landscape of bacteria. *Genome Biol* **10**: 233.
- Velculescu, V.E., Zhang, L., Zhou, W., Vogelstein, J., Basrai, M.A., Bassett, D.E., Jr., *et al.* (1997) Characterization of the Yeast Transcriptome. *Cell* **88**: 243–251.
- Vergara-Irigaray, M., Maira-Litran, T., Merino, N., Pier, G.B., Penades, J.R., and Lasa, I. (2008) Wall teichoic acids are dispensable for anchoring the PNAG exopolysaccharide to the *Staphylococcus aureus* cell surface. *Microbiology* **154**: 865–877.
- Vergara-Irigaray, M., Valle, J., Merino, N., Latasa, C., García, B., Ruiz de los Mozos, I., *et al.* (2009) Relevant role of fibronectin-binding proteins in *Staphylococcus aureus* biofilm-associated foreign-body infections. *Infect Immun* **77**: 3978–3991.
- Vincent, H.A., and Deutscher, M.P. (2006) Substrate recognition and catalysis by the exoribonuclease RNase R. *J Biol Chem* **281**: 29769–29775.
- Vlasova, I.A., Tahoe, N.M., Fan, D., Larsson, O., Rattenbacher, B., SternJohn, J.R., *et al.* (2008) Conserved GU-Rich Elements Mediate mRNA Decay by Binding to CUG-Binding Protein 1. *Molecular Cell* **29**: 263–270.
- Vogel, C., and Marcotte, E.M. (2012) Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13**: 227–232.

- Vuong, C., Kocianova, S., Voyich, J.M., Yao, Y., Fischer, E.R., DeLeo, F.R., and Otto, M. (2004) A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion, and virulence. *J Biol Chem* **279**: 54881–54886.
- Wachi, M., Umitsuki, G., Shimizu, M., and Takada, A. (1999) *Escherichia coli* *cafA* Gene Encodes a Novel RNase, Designated as RNase G, Involved in Processing of the 5' End of 16S rRNA. *Biochemical and Biophysical Research Communications* **259**: 483–488.
- Wagner, C., Palacios, I., Jaeger, L., St Johnston, D., Ehresmann, B., Ehresmann, C., and Brunel, C. (2001) Dimerization of the 3'UTR of *bicoid* mRNA involves a two-step mechanism. *Journal of Molecular Biology* **313**: 511–524.
- Waldminghaus, T., Heidrich, N., Brantl, S., and Narberhaus, F. (2007) FourU: a novel type of RNA thermometer in *Salmonella*. *Molecular Microbiology* **65**: 413–424.
- Wang, Z., Gerstein, M., and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10**: 57–63.
- Wassarman, K.M., Repoila, F., Rosenow, C., Storz, G., and Gottesman, S. (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes & Development* **15**: 1637–1651.
- Waters, L.S., and Storz, G. (2009) Regulatory RNAs in Bacteria. *Cell* **136**: 615–628.
- Wertheim, H., Melles, D., VOS, M., Vanleeuwen, W., Vanbelkum, A., Verbrugh, H., and Nouwen, J. (2005) The role of nasal carriage in infections. *The Lancet Infectious Diseases* **5**: 751–762.
- Westhof, E., Masquida, B., and Jossinet, F. (2011) Predicting and modeling RNA architecture. *Cold Spring Harbor Perspectives in Biology* **3**: a003632
- Wilson, K.S., and Hippel, von, P.H. (1995) Transcription termination at intrinsic terminators: the role of the RNA hairpin. *PNAS* **92**: 8793–8797
- Winkler, W.C., and Breaker, R.R. (2005) Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* **59**: 487–517.
- Woese, C.R. (1967) *The genetic code: The molecular basis for genetic expression*. Harper & Row New York.
- Woese, C.R. (1987) Bacterial evolution. *Microbiol Rev* **51**: 221–271.
- Worrall, J.A.R., and Luisi, B.F. (2007) Information available at cut rates:

structure and mechanism of ribonucleases. *Current Opinion in Structural Biology* **17**: 128–137.

Wurtzel, O., Sesto, N., Mellin, J.R., Karunker, I., Edelheit, S., Bécavin, C., *et al.* (2012) Comparative transcriptomics of pathogenic and non-pathogenic *Listeria* species. *Mol Syst Biol* **8**: 583.

Wurtzel, O., Yoder-Himes, D.R., Han, K., Dandekar, A.A., Edelheit, S., Greenberg, E.P., *et al.* (2012) The Single-Nucleotide Resolution Transcriptome of *Pseudomonas aeruginosa* Grown in Body Temperature. *PLoS Pathog* **8**: e1002945.

Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., *et al.* (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**: 842–846.

Yan, M., Pamp, S.J., Fukuyama, J., Hwang, P.H., Cho, D.-Y., Holmes, S., and Relman, D.A. (2013) Nasal Microenvironments and Interspecific Interactions Influence Nasal Microbiota Complexity and *S. aureus* Carriage. *Cell Host & Microbe* **14**: 631–640.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Research* **30**: e15

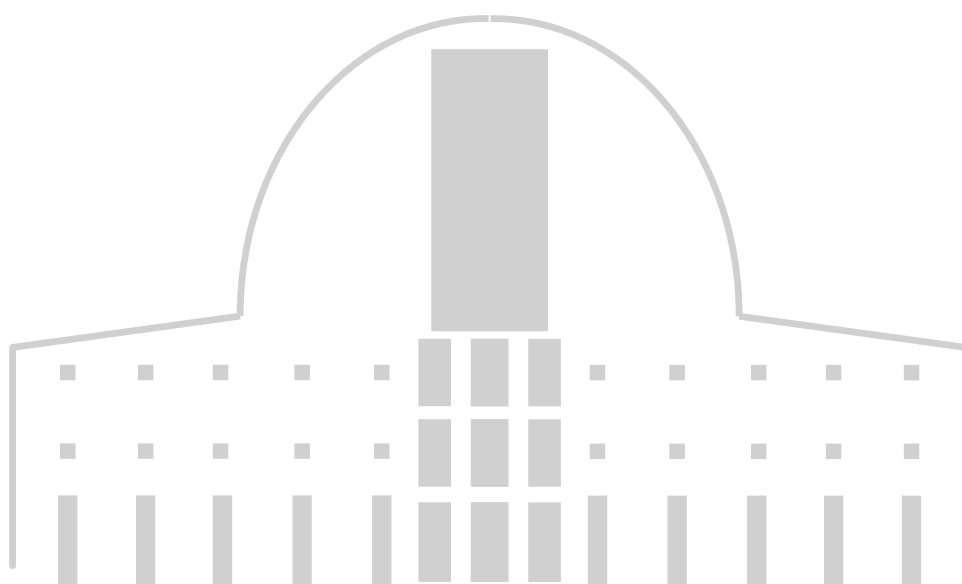
Yazaki, J., Gregory, B.D., and Ecker, J.R. (2007) Mapping the genome landscape using tiling array technology. *Current Opinion in Plant Biology* **10**: 534–542.

Zhu, L.Q., Gangopadhyay, T., Padmanabha, K.P., and Deutscher, M.P. (1990) *Escherichia Coli* Rna Gene Encoding Rnase I Cloning, Overexpression, Subcellular-Distribution of the Enzyme, and Use of an Rna Deletion to Identify Additional Rnases. *Journal of Bacteriology* **172**: 3146–3151.

Zilka, A., Garlapati, S., Dahan, E., Yaolsky, V., and Shapira, M. (2001) Developmental Regulation of Heat Shock Protein 83 in *Leishmania*. *J Biol Chem*. **64**: 95-110

Zuker, M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**: 3406–3415.

Zuo, Y., Vincent, H.A., Zhang, J., Wang, Y., Deutscher, M.P., and Malhotra, A. (2006) Structural basis for processivity and single-strand specificity of RNase II. *Molecular Cell* **24**: 149–156.



Universidad Pública de Navarra
Departamento de Producción Agraria
Instituto de Agrobiotecnología